A mixed-step algorithm for the approximation of the stationary regime of a diffusion

Gilles Pagès, Fabien Panloup

A mixed-step algorithm for the approximation of the stationary regime of a diffusion

Gilles Pagès[*], Fabien Panloup[†]

March 8, 2013

Abstract

In some recent papers, some procedures based on some weighted empirical measures related to decreasing-step Euler schemes have been investigated to approximate the stationary regime of a diffusion (possibly with jumps) for a class of functionals of the process. This method is efficient but needs the computation of the function at each step. To reduce the complexity of the procedure (especially for functionals), we propose in this paper to study a new scheme, called *mixed-step scheme* where we only keep some regularly time-spaced values of the Euler scheme. Our main result is that, when the coefficients of the diffusion are smooth enough, this alternative does not change the order of the rate of convergence of the procedure. We also investigate a Richardson-Romberg method to speed up the convergence and show that the variance of the original algorithm can be preserved under a uniqueness assumption for the invariant distribution of the "duplicated" diffusion, condition which is extensively discussed in the paper. Finally, we end by giving some sufficient "asymptotic confluence" conditions for the existence of a smooth solution to a discrete version of the associated Poisson equation, condition which is required to ensure the rate of convergence results.

Keywords: stochastic differential equation; stationary process; steady regime; ergodic diffusion; Central Limit Theorem; Euler scheme; Poisson equation ; Richardson-Romberg extrapolation.

AMS classification (2000): 60G10, 60J60, 65C05, 65D15, 60F05.

1 Introduction

In a series of papers ([11, 13, 19, 20, 17, 18]) going back to [10], have been investigated the properties of an Euler scheme with decreasing step as a tool for the numerical approximation of the [steady/stationary] regime of a diffusion, possibly with jumps, satisfying some [stability/mean reverting] conditions. The purpose of the present paper is to propose and investigate a variant of the original procedure sharing the similar properties in terms of convergence and rate but with a lower complexity, especially in its functional form, *i.e.* when trying to compute the expectation of a functional of the process (over a finite time interval $[0,T]$) with respect to the stationary distribution of the process.

[*]Laboratoire de Probabilités et Modèles Aléatoires, UMR 7599, UPMC, Case 188, 4 pl. Jussieu, F-75252 Paris Cedex 5, France, E-mail: gilles.pages@upmc.fr

[†]Institut de Mathématiques de Toulouse, Université Paul Sabatier & INSA Toulouse, 135, av. de Rangueil, F-31077 Toulouse Cedex 4, France, E-mail: fabien.panloup@math.univ-toulouse.fr

In this paper we will focus on the case of Brownian diffusions. We consider an $\mathbb{R}^d$-valued diffusion process (X_t) solution to

$$(SDE) \ \equiv dX_t = b(X_t)dt + \sigma(X_t)dW_t, \tag{1.1}$$

where $(W_t)_{t\geq 0}$ is a q-dimensional standard Brownian motion (SBM) and the coefficients b and σ are *Lipschitz continuous* functions from $\mathbb{R}^d$ to $\mathbb{R}^d$ and $\mathbb{M}_{d,q}$ respectively (where $\mathbb{M}_{d,q}$ denotes the set of $d\times q$-matrices). Under these assumptions, strong existence and uniqueness hold for the SDE starting from any $\mathbb{R}^d$-valued r.v. independent of W and $(X_t)_{t\geq 0}$ is a homogeneous Markov process with semi-group $(P_t)_{t\geq 0}$. We will denote by $\mathbb{P}_\mu$, the distribution of the whole process $(X_t)_{t\geq 0}$ (supported by the set $\mathcal{C}(\mathbb{R}_+,\mathbb{R}^d)$ of continuous functions from $\mathbb{R}_+$ to $\mathbb{R}^d$) when starting from X_0 with distribution μ. We also assume throughout the paper that $(X_t)_{t\geq 0}$ has a unique invariant distribution ν. Except in the one-dimensional case, ν is usually not explicit and the numerical computation of ν or $\mathbb{P}_\nu$ (which, in particular, is fundamental, to estimate the asymptotic behavior of ergodic processes) then requires some specific numerical methods.

Let us briefly describe the discrete and continuous time Euler schemes with decreasing step resulting from the time discretization of the diffusion $(X_t)_{t\geq 0}$. First we introduce a non-decreasing sequence $(\Gamma_n)_{n\geq 1}$ sequence of discretization times starting from $\Gamma_0 = 0$ and we assume that the step sequence defined as its increments by $\gamma_n := \Gamma_n - \Gamma_{n-1}$, $n \geq 1$, is *nonincreasing* and satisfies

$$\lim_{n\to+\infty} \gamma_n = 0 \quad \text{and} \quad \Gamma_n = \sum_{k=1}^n \gamma_k \xrightarrow{n\to+\infty} +\infty. \tag{1.2}$$

The discrete time Euler scheme $(\bar{X}_{\Gamma_n})_{n\geq 0}$ (with Brownian increments) is recursively defined at discretization times Γ_n by $\bar{X}_0 = x_0$ and

$$\bar{X}_{\Gamma_{n+1}} = \bar{X}_{\Gamma_n} + \gamma_{n+1} b(\bar{X}_{\Gamma_n}) + \sigma(\bar{X}_{\Gamma_{n+1}})(W_{\Gamma_{n+1}} - W_{\Gamma_n}). \tag{1.3}$$

If we introduce the notation

$$\underline{t} = \Gamma_{N(t)} \quad \text{with} \quad N(t) = \min\{n \geq 0, \Gamma_{n+1} > t\} \tag{1.4}$$

so that $\underline{t} = \Gamma_k$ if and only if $t \in [\Gamma_k, \Gamma_{k+1})$, the stepwise constant Euler scheme also reads

$$\forall\, t \in \mathbb{R}_+, \qquad \bar{X}_t = \bar{X}_{\underline{t}}.$$

The idea at the origin of [10] was to make the guess, mimicking the pointwise ergodic theorem, that the weighted empirical measure

$$\nu_n(\omega, dx) = \frac{1}{\Gamma_n} \sum_{k=1}^n \gamma_k \delta_{\bar{X}_{\Gamma_{k-1}}(\omega)}(dx) \tag{1.5}$$

or more generally

$$\nu_n^\eta(\omega, dx) = \frac{1}{H_n} \sum_{k=1}^n \eta_k \delta_{\bar{X}_{\Gamma_{k-1}}(\omega)}(dx)$$

where $(\eta_n)_{n\geq 1}$is a sequence of positive weights such that $H_n = \eta_1 + \cdots + \eta_n \to \infty$ would weakly converge $a.s.$ to the invariant distribution ν under appropriate assumptions on both step and weight sequences $(\gamma_n)_{n\geq 1}$ and $(\eta_n)_{n\geq 1}$ and on the drift b and σ. Typically, some mean-reverting assumptions stated through the existence of an "essentially quadratic" twice differentiable coercive function $V : \mathbb{R}^d \to (0;+\infty)$ for which there exists $a \in (0,1)$ such that

$$\mathcal{A}V \le \beta - \alpha V^a$$

where $\mathcal{A}$ stands for the infinitesimal generator of the diffusion. In fact, as far as the weighted empirical measures of the Euler scheme are concerned, slightly more stringent conditions are required.

The $a.s$ weak convergence of ν_n^η toward ν as well as its rate of convergence (CLT or $a.s$ depending on the step rate of decay) has been extensively investigated in the above cited references, including much more general setting than Brownian diffusions. Basically when the step goes to 0 fast enough the empirical measure behaves like the empirical measure of the diffusion itself and satisfies the standard CLT (see *e.g.* [2] and [18] for results on the diffusion itself). When the convergence of (γ_n) to 0 is too slow, the whole procedure is slowed down and satisfies an $a.s.$ (or at least in probability) rate of convergence property. The critical rate is obtained with $\gamma_n = n^{-\frac{1}{3}}$ with a biased CLT at rate $n^{\frac{1}{3}}$.

Then, motivated by problems arising in Finance, pricing of exotic derivatives (Asian and barrier options) in stationary (stochastic) volatility models, this approach has been extended to functionals and led to investigate (when $\eta_n = \gamma_n$) the behaviour of the empirical measures $(\nu^{(n),T}(\bar{X}(\omega), d\alpha))$ or $(\nu^{(n),T}(\xi(\omega), d\alpha))$ where $\bar{X}$ and ξ denote respectively the stepwise constant and continuous time Euler schemes and, for a càdlàg function $Y : \mathbb{R}_+ \to \mathbb{R}^d$, $\nu^{(n),T}(Y, d\alpha)$ is defined on the set of probability measures on the Skorokhod set $\mathbb{D}([0,T], \mathbb{R}^d)$, $T > 0$, denoting a fixed horizon,

$$\nu^{(n),T}(Y, d\alpha) = \frac{1}{\Gamma_n} \sum_{k=1}^{n} \gamma_k \delta_{Y^{(\Gamma_k),T}}(d\alpha), \; n \ge 1$$

where, by definition, for every càdlàg function $\alpha \in \mathbb{D}(\mathbb{R}_+, \mathbb{R}^d)$

$$\alpha^{(t),T}(s) = \alpha(t+s), \; s \in [0,T].$$

What we call continuous time Euler scheme here (also sometimes known in the literature as the *genuine* Euler scheme) is the process $(\xi_t)_{t\ge 0}$ defined as interpolation of the stepwise constant Euler scheme, each driving term (time and Brownian motion) being interpolated in its own scale:

$$\forall n \in \mathbb{N}, \quad \forall t \in [\Gamma_n, \Gamma_{n+1}), \quad \xi_t = \bar{X}_{\Gamma_n} + (t - \Gamma_n) b(\bar{X}_{\Gamma_n}) + \sigma(\bar{X}_{\Gamma_n})(W_t - W_{\Gamma_{n+1}}). \tag{1.6}$$

A convenient and synthetic form for the genuine Euler scheme is to write it as an Itô process satisfying the following pseudo-diffusion equation

$$\xi_t = X_0 + \int_0^t b(\xi_{\underline{s}}) ds + \int_0^t \sigma(\xi_{\underline{s}}) dW_s. \tag{1.7}$$

Such an approximation looks more accurate than the former one, especially when dealing with functionals of the process, as it has been emphasized – in the constant step framework – in the literature on several problems related to the Monte Carlo estimation of ($a.s.$ continuous) functionals of a diffusion (with a finite horizon) (see *e.g.* [4], Chapter 5). This follows from the classical fact that the L^p-convergence rate of this scheme for the sup norm is $\sqrt{\gamma}$ instead of $\sqrt{\gamma}\log(1/\gamma)$ for its stepwise constant counterpart (where γ stands for the step). On the other hand, the simulation of a functional of $(\xi_t)_{t \in [\tau, \tau+T]}$ is deeply connected with the simulation of functionals of the Brownian bridge so that it is only possible for few specific such functionals (like running maxima, running means,…).

These empirical measures on path spaces have in turn been extensively investigated in [17, 18]. However, their practical implementation suffers in practice from a significant drawback: one cannot prevent the complexity of the computation of $F(\bar{X}^{(\Gamma_n),T})$ to go to infinity as n grows since more and more iterations of the Euler schemes are needed to "cover" a laps of time equal to T. In practice, within the usual range of accuracy requested, no storing difficulty has been encountered. Furthermore in many standard situations where the functional F only involves running maxima or time integrals some recursive procedures make possible not to store the whole path of "length" T. It remains that the computational cost remains high.

As concerns the computation of the (marginal) invariant distribution, this led us naturally to consider a mixed step procedure where the decreasing step Euler scheme would be treated as if it were a constant step Euler scheme with step $T>0$. This means, to implement and investigate the convergence of the sequence of empirical measures

$$\bar{\mu}_n(\omega,dx)=\frac{1}{n}\sum_{k=0}^{n-1}\delta_{\xi_{kT}(\omega)}(dx),\ n\geq 1,$$

having in mind that the genuine Euler scheme can always be simulated at a locally finite number of additional times. Our aim is to prove that its empirical measure $a.s.$ weakly converges toward ν under the same assumptions as ν_n with the same rate structure (however an additional uniqueness assumption of ν with respect to P_T is requested as one could expect). Of course the asymptotic variance/bias will differ: as for ν_n it was related to the continuous Poisson equation $f-\nu(f)=-\mathcal{A}\varphi$ while in this new framework, it will be related to the Poisson equation involving the pseudo-infinitesimal generator $f-\nu(f)=\frac{Id-P_T}{T}(\varphi)$. Furthermore, the fact that we only "sample" the scheme at times nT will imply a control of the discretization error on intervals of length T. More precisely, we will rely here on an extension of the classical weak error results "à la Talay-Tubaro" (see [24]) in our decreasing step framework, also known as the PDE method. This part is self-contained since we will directly establish by probabilistic methods the spatial regularity of the objects to be plugged in the PDE.

We will also investigate the functional counterpart of $\bar{\mu}_n(\omega,dx)$, denoted $\bar{\mu}^{(n)}(\omega,d\alpha)$, and defined on (the Borel σ-field of) $\mathcal{C}(\mathbb{R}_+,\mathbb{R}^d)$ by

$$\forall n\in\mathbb{N},\quad \bar{\mu}^{(n)}(\omega,d\alpha)=\frac{1}{n}\sum_{k=0}^{n-1}\delta_{\xi^{(kT)}(\omega)}(d\alpha)$$

where for any $t>0$, $(\xi_s^{(t)})_{s\geq 0}$ is defined by $\xi_s^{(t)}=\xi_{t+s}$.

The second aim of the paper is to illustrate how to implement a specific Romberg extrapolation method in this framework, in order to "erase" partially the slowing effect due to the time discretization of the underlying diffusion. We will show that, provided the decreasing step sequence and the constant pseudo-step T are consistent, this allows for a significant speeding-up of the procedure (as well as an extension of the range of the CLT as a rule for the convergence rate). Furthermore in order to control the asymptotic variance it seems natural to consider consistent Brownian increments for the two schemes, as emphasized in [16] in a constant step framework. However, we will see that the situation is more involved when dealing with long run behaviour. In fact, the best strategy depends on the long run behaviour of the duplicated diffusion system

$$\begin{cases} dX_t=b(X_t)dt+\sigma(X_t)dW_t\\ dX_t^{(\theta)}=b(X_t^{(\theta)})dt+\sigma(X_t^{(\theta)})(\theta dW_t+\sqrt{1-\theta^2}d\widetilde{W}_t)\end{cases}$$

where $\widetilde{W}$ and W are independent and $\theta \in [-1,1]$. We will see that the strategy is actually optimal with $\theta = 1$ (*i.e.* with consistent Brownian increments) but under an additional assumption: we will need to assume that the invariant measure of the duplicated diffusion (with $\rho = 1$) is unique (see Sections 3.4 and 3.5 for more details).

The paper is organized as follows. Section 2 is devoted to some additional notations and to some background on some existing results which are required for our study. In Section 3 are stated the main results of the paper including the functional case, the Romberg extrapolation and a brief review of some conditions of uniqueness of the invariant distribution of the above duplicated system. Section 4, 5 and 6 are devoted to the proof of the main theorems, including new results on the weak error. Finally, in section 7, we give some asymptotic confluence conditions which ensure the existence of regular solutions to Poisson-type equations (see condition $(\mathcal{P}_{\mathbf{k},\mathbf{T}})$).

2 Background and Notations

2.1 Notations

$\triangleright$ $\langle x, y\rangle = \sum_i x_i y_i$ will denote the canonical inner product and $|x| = \sqrt{\langle x, x\rangle}$ will denote Euclidean norm of a vector $x \in \mathbb{R}^d$. For every $k \in \mathbb{N}$, we denote by $\mathcal{C}^k(\mathbb{R}^d)$ the set of functions $f : \mathbb{R}^d \to \mathbb{R}$ whose derivatives up to order k are continuous and by $\mathcal{C}^{k,\alpha}(\mathbb{R}^d)$ ($\alpha \in (0,1]$) the subset of $\mathcal{C}^k(\mathbb{R}^d)$ such that for every $k_1, \ldots, k_d \in \mathbb{N}$ such that $k_1 + \ldots + k_d = k$, $\partial^k_{x_1^{k_1}\cdots x_d^{k_d}}$ is a (locally) α-Hölder continuous function.

$\triangleright$ Let $A = [a_{ij}] \in \mathbb{M}_{d,q}$ be an $\mathbb{R}$-valued matrix with d rows and q columns. A^* will denote the transpose of A, $\mathrm{Tr}(A) = \sum_i a_{ii}$ its trace and $\|A\| := \sqrt{\mathrm{Tr}(AA^*)} = (\sum_{ij} a_{ij}^2)^{\frac{1}{2}}$. If $d = q$, one writes $A\, x^{\otimes 2}$ for $x^* A x$.

$\triangleright$ We denote by $\mathcal{A}$ the infinitesimal generator of $(X_t)_{t\geq 0}$ defined for every $f \in \mathcal{C}^2(\mathbb{R}^d)$ by

$$\forall x \in \mathbb{R}^d, \quad \mathcal{A}f(x) = \langle \nabla f(x), b(x)\rangle + \frac{1}{2}\mathrm{Tr}(\sigma^*(x) D^2 f(x) \sigma(x))$$

where for every $x \in \mathbb{R}^d$, $D^2 f(x)$ denotes the Hessian matrix of f defined by $(D^2 f)_{ij}(x) = \partial^2_{x_i x_j} f(x)$, $i, j \in \{1, \ldots, d\}$. We also denote by $(\mathcal{A}^{(k)})_{k\geq 1}$ the sequence of operators recursively defined by $\mathcal{A}^{(1)} = \mathcal{A}$ and

$$\forall f \in \mathcal{C}^{2k}(\mathbb{R}^d), \quad \mathcal{A}^{(k+1)} f = \mathcal{A}(\mathcal{A}^{(k)} f)$$

as long as b and σ are smooth enough. As concerns the discretized process $(\xi_t)_{t\geq 0}$, we also introduce the following associated operators $\bar{\mathcal{A}}^{(k)}$ recursively defined for $f \in \mathcal{C}^{2k}(\mathbb{R}^d)$, for every $\underline{x} \in \mathbb{R}^d$ and $k \in \mathbb{N}$ by,

$$\bar{\mathcal{A}} f(.,\underline{x}) = \langle \nabla f(.), b(\underline{x})\rangle + \frac{1}{2}\mathrm{Tr}(\sigma^*(\underline{x}) D^2 f(.) \sigma(\underline{x})) \quad \text{and,} \quad \bar{\mathcal{A}}^{(k+1)} f(.,\underline{x}) = \bar{\mathcal{A}}(\bar{\mathcal{A}}^{(k)} f(.,\underline{x})).$$

$\triangleright$ We denote by $(\mathcal{F}_t)_{t\geq 0}$ the usual augmentation of $(\sigma(W_s, 0 \leq s \leq t))_{t\geq 0}$ by $\mathbb{P}$-negligible sets. For $t \geq 0$, we also set $\mathbb{E}_t[\,.\,] = \mathbb{E}[\,.\,|\mathcal{F}_t]$.

$\triangleright$ When necessary, we will adopt the more precise notations $\bar{X}^{x,(h_n)}$ or $\xi^{x,(h_n)}$ for the stepwise constant or genuine continuous-time Euler schemes respectively, in order to specify that these processes start from $x \in \mathbb{R}^d$ at time 0 and that their discretization step sequence is $(h_n)_{n\geq 1}$.

2.2 Convergence results

In this part, we recall some $a.s.$-convergence results for $(\bar{\mu}_n(\omega,dx))_{n\geq 1}$ to the invariant distribution. To this end, we denote by $\mathcal{EQ}(\mathbb{R}^d)$ the set of positive $\mathcal{C}^2$-functions $V:\mathbb{R}^d\to\mathbb{R}$ such that there exists $\rho>0$ such that

$$\liminf_{|x|\to+\infty}|x|^{-\rho}V(x)>0,\quad |\nabla V|^2\leq CV\quad\text{and}\quad D^2V \text{ is bounded.}$$

Note that $\lim_{|x|\to+\infty}|x|^{-\rho}V(x)>0$ implies that $\lim_{|x|\to+\infty}V(x)=+\infty$ which is a sufficient assumption for a part of the results. The interest of this slightly more restrictive assumption is to say that a function g has polynomial growth if and only if there exists $r>0$ such that $|g|\leq CV^r$.

Then, for any symmetric $d\times d$ matrix S, set $\lambda_S^+:=\max(0,\lambda_1,\ldots,\lambda_d)$ where $\lambda_1,\ldots,\lambda_d$ denote the eigenvalues of S. Let $a\in(0,1]$ and $p\in[1,+\infty)$. We introduce the following mean-reverting assumption *with intensity* a:

$(\mathbf{S_{a,p}})$: There exists a function $V\in\mathcal{EQ}(\mathbb{R}^d)$ such that:

- (i) $\exists C_a>0$ such that $|b|^2+\mathrm{Tr}(\sigma\sigma^*)\leq C_aV^a$.
- (ii) There exist $\beta\in\mathbb{R}$ and $\rho>0$ such that $\langle\nabla V,b\rangle+\lambda_p\mathrm{Tr}(\sigma\sigma^*)\leq\beta-\rho V^a$,

where $\lambda_p:=\dfrac{1}{2}\sup_{x\in\mathbb{R}^d}\lambda^+_{D^2V(x)+(p-1)\frac{\nabla V\otimes\nabla V}{V}}$. The function V is then called a Lyapunov function for the diffusion $(X_t)_{t\geq0}$. If $(\mathbf{S_{a,p}})$ holds for every $p\in[1,\infty)$, we denote it by $(\mathbf{S_{a,\infty}})$. Note that $(\mathbf{S_{a,\infty}})$ holds if $|b|^2\leq C_aV^a$, $\langle\nabla V,b\rangle\leq\beta-\rho V^a$ and $\mathrm{Tr}(\sigma\sigma^*)=o(V^a)$.

We finally introduce a uniqueness assumption for the invariant distribution.

$(\mathbf{S^\nu_T})$: ν is an invariant distribution for $(P_t)_{t\geq0}$ and the unique one for P_T.

Then, the following result follows from Lemma 3.3 from [18]:

Proposition 2.1. *Assume* $(\mathbf{S_{a,p}})$ *holds with* $p>2$ *and* $a in(0,1]$. *Then,*

(i) *The sequence of empirical measures* $(\bar{\mu}_n(\omega,dx))_{n\geq1}$ *satisfies*

$$\sup_{n\geq1}\bar{\mu}_n(\omega,V^{\frac{p}{2}+a-1})<+\infty\quad a.s. \tag{2.8}$$

In particular, $(\bar{\mu}_n(\omega,dx))_{n\geq1}$ *is a.s. tight.*

(ii) *Assume* $(\mathbf{S^\nu_T})$. *Then, a.s., for every continuous function* f *such that* $f(x)=o(V^{\frac{p}{2}+a-1}(x))$ *as* $|x|\to+\infty$. *Then,* $\bar{\mu}_n(\omega,f)\xrightarrow{n\to+\infty}\nu(f)$ *a.s.*

2.3 Smoothness and growth of solutions to parabolic PDE's

Let $T\in(0,\infty)$ and let $g\in\mathcal{C}^2(\mathbb{R}^d,\mathbb{R})$. Set for every $(t,x)\in[0,T]\times\mathbb{R}^d$, $u_g(t,x):=\mathbb{E}[g(X^x_{T-t})]$ where X^x denotes the solution to (1.1) starting from $x\in\mathbb{R}^d$. Owing to Itô's formula and to the commutation property between $\mathcal{A}$ and P_{T-t}, u_g is a solution to the parabolic PDE:

$$\begin{cases}\partial_tu_g(t,x)+\mathcal{A}u_g(t,x)=0 & (t,x)\in[0,T]\times\mathbb{R}^d,\\ u_g(T,x)=g(x), & x\in\mathbb{R}^d.\end{cases} \tag{2.9}$$

In fact, the main point is to obtain smoothness (and growth control) properties for u_g. This is the purpose of the following proposition.

Proposition 2.2. *Let m be a positive integer. Assume that b and σ are $\mathcal{C}^{m,\alpha}$-functions ($\alpha \in (0,1]$) on $\mathbb{R}^d$ with bounded derivatives. Assume that g is a $\mathcal{C}^m$-function on $\mathbb{R}^d$ such that g and its derivatives of order $l \leq m$ have polynomial growth. Then, $x \mapsto u_g(t,x)$ is a $\mathcal{C}^m$-function on $\mathbb{R}^d$ and $t \mapsto u_g(t,x)$ is a $\mathcal{C}^{\lfloor m/2 \rfloor}$-function on $\mathbb{R}_+$. Furthermore, there exists $r > 0$ such that $\forall\, t \in [0,T]$,*

$$|u_g(t,x)| \leq C_T(1+|x|^r) \quad \text{and} \quad |\partial^{\ell}_{x_1^{\beta_1}\ldots x_d^{\beta_d}} u_g(t,x)| \leq C_T(1+|x|^r), \tag{2.10}$$

for every $\ell \in \{1,\ldots,m\}$ and $\beta_1,\ldots,\beta_d \in \mathbb{N}^d$ such that $\beta_1+\ldots+\beta_d = \ell$. Finally, in the particular case where g and its derivatives are bounded, u and its derivatives are also bounded.

We provide references and a direct self-contained probabilistic proof in the appendix.

3 Main results

3.1 Marginal case

Let $f : \mathbb{R}^d \to \mathbb{R}$ be a continuous function. Before stating the first result of this paper about the rate of convergence of $(\bar{\mu}_n(\omega, f))_{n \geq 1}$, we need to introduce the following "co-boundary" assumption on f (where m is an integer):

$(\mathcal{P}_{\mathbf{m},\mathbf{T}})$: There exists a $\mathcal{C}^m$-function $g_{T} : \mathbb{R}^d \to \mathbb{R}$ such that

$$f(x) - \nu(f) = \frac{g_T(x) - P_T g_T(x)}{T}, \quad x \in \mathbb{R}^d. \tag{3.11}$$

Furthermore, if g_T and its derivatives are bounded functions, we will say that f satisfies $(\mathcal{P}^{\mathbf{b}}_{\mathbf{m},\mathbf{T}})$. Likewise, if g_T and its derivatives have polynomial growth, we will denote the assumption by $(\mathcal{P}^{\mathrm{pol}}_{\mathbf{m},\mathbf{T}})$.

Remark 3.1. Assumption $(\mathcal{P}_{\mathbf{m},\mathbf{T}})$ can be viewed as the existence of a regular solution to a discrete version of the Poisson equation $f - \nu(f) = -\mathcal{A}g$. In Section 7, we give some criteria which ensure such condition when the diffusion is asymptotically confluent (see Proposition 7.8). We refer to [21, 22, 23] for results on the Poisson equation itself in an elliptic setting. Since T is a fixed positive real number throughout the paper, we will usually write g instead of g_T in order to alleviate the notations.

Theorem 3.1. *Assume $(\mathbf{S_{a,p}})$ holds with $a \in (0,1]$ and $p \in (2,+\infty]$. Assume $(\mathbf{S}^{\nu}_{\mathbf{T}})$. Let $m \in \mathbb{N}$ and let $\alpha \in (0,1]$ such that b and σ are $\mathcal{C}^{m,\alpha}$-functions on $\mathbb{R}^d$ with bounded derivatives. Let $f : \mathbb{R}^d \to \mathbb{R}$ be a Borel function satisfying $(\mathcal{P}^{\mathbf{b}}_{\mathbf{m},\mathbf{T}})$ if $p < +\infty$ or $(\mathcal{P}^{\mathrm{pol}}_{\mathbf{m},\mathbf{T}})$ if $p = +\infty$. Then,*

(i) If $m = 4$ and $\frac{1}{\sqrt{n}} \sum_{k=1}^n \gamma_{N(kT)} \xrightarrow{n \to +\infty} 0$,

$$\sqrt{nT}\left(\frac{1}{n}\sum_{k=1}^n f(\xi_{(k-1)T}) - \nu(f)\right) \overset{n \to +\infty}{\Longrightarrow} \mathcal{N}(0; \hat{\sigma}^2_T)$$

with $\hat{\sigma}^2_T = \dfrac{1}{T}\displaystyle\int \left(g^2_T(x) - (P_T g_T(x))^2\right)\nu(dx)$.

(ii) If $m = 5$ and $\frac{1}{\sqrt{n}} \sum_{k=1}^n \gamma_{N(kT)} \xrightarrow{n \to +\infty} +\infty$ and $\gamma_{N(nT)} - \gamma_{N((n+1)T)} = o(\gamma_{N(nT)})$ as $n \to +\infty$,

$$\frac{n}{\sum_{k=1}^n \gamma_{N(kT)}}\left(\frac{1}{n}\sum_{k=1}^n f(\xi_{(k-1)T}) - \nu(f)\right) \xrightarrow{\mathbb{P}} m_T = \frac{1}{2T}\int_{\mathbb{R}^d}\int_0^T \mathbb{E}[\Phi_g(s, X^x_s)]ds\nu(dx)$$

as $n \to +\infty$, where "$\xrightarrow{\mathbb{P}}$" denotes the convergence in probability and

$$\Phi_g(t,x) = \langle \mathcal{A}(\partial_x u_g)(t,x) - \partial_x(\mathcal{A}u_g)(t,x), b(x)\rangle + \left(\mathcal{A}(\partial^2_{x^2} u_g)(t,x) - \partial^2_{x^2}(\mathcal{A}u_g)(t,x)\right)\sigma(x)^{\otimes 2} \quad (3.12)$$

where $u(t,x) = \mathbb{E}[g(X^x_{T-t})]$ and $\partial_x u_g$ and $\partial^2_{x^2} u_g$ denote the gradient and the Hessian matrix of $x \mapsto u_g(t,x)$ respectively.

(iii) If $m = 5$ and $\sqrt{\frac{T}{n}}\sum_{k=1}^n \gamma_{N(kT)} \xrightarrow{n\to+\infty} \beta_0 \in (0,+\infty)$ and $\gamma_{N(nT)} - \gamma_{N((n+1)T)} = o(\gamma_{N(nT)})$ as $n \to +\infty$,

$$\sqrt{nT}\left(\frac{1}{n}\sum_{k=1}^n f(\xi_{(k-1)T}) - \nu(f)\right) \stackrel{n\to+\infty}{\Longrightarrow} \mathcal{N}(\beta_0 m_{_T}; \hat{\sigma}^2_T).$$

Remark 3.2. Owing to the stationarity property, one can check that $\hat{\sigma}^2_T$ can also be written

$$\hat{\sigma}^2_T = \frac{1}{T}\int \mathbb{E}\left[\left(g_{_T}(X^x_T) - \mathbb{E}[g_{_T}(X^x_T)]\right)^2\right]\nu(dx). \quad (3.13)$$

3.2 Comparison with the original procedure

Throughout this section, we compare $(\bar{\mu}_n(\omega,dx))_{n\geq 1}$ with the original procedure $(\nu_n(\omega,dx))_{n\geq1}$ defined by (1.5).

Assume that $\gamma_k = \gamma_1 k^{-\rho}$ with $\rho \in (0,1)$. Then, $\Gamma_n = \gamma_1 n^{1-\rho} + O(1)$ and $N(t) = \gamma_1^{-\frac{1}{1-\rho}} t^{\frac{1}{1-\rho}} + O(1)$. It follows that

$$\gamma_{N(nT)} = \gamma_1^{\frac{1}{1-\rho}} T^{-\frac{\rho}{1-\rho}} n^{-\frac{\rho}{1-\rho}} + O(n^{-\frac{1+\rho}{1-\rho}})$$

and thus, that

$$\frac{1}{\sqrt{n}}\sum_{k=1}^n \gamma_{N(kT)} \xrightarrow{n\to+\infty} 0 \iff \rho > \frac{1}{3}.$$

Likewise, since $\gamma_{N(nT)} - \gamma_{N((n+1)T)} = O(n^{-\frac{1}{1-\rho}})$, $\gamma_{N(nT)} - \gamma_{N((n+1)T)} = o(\gamma_{N(nT)})$ for every $\rho \in (0,1)$. At time $t_n = nT$, we deduce that the error is of order $(t_n)^{-\frac{1}{2}}$ if $\rho > 1/3$ and of order $(t_n)^{-\frac{\rho}{1-\rho}}$ is $\rho \leq 1/3$. Now, from a numerical point of view, we must compute the error in terms of the number N of discretization times. The error is then of order

$$\begin{cases} \Gamma_N^{-\frac{1}{2}} \propto N^{-\frac{1-\rho}{2}} & \text{if } \rho \in (1/3,1) \\ \Gamma_N^{-\frac{\rho}{1-\rho}} \propto N^{-\rho} & \text{if } \rho \in (0,1/3]. \end{cases} \quad (3.14)$$

This means that we exactly retrieve the error orders obtained for the original procedure $(\nu_N(\omega,dx))_{N\geq1}$ (see [10]). Thus, as announced in the abstract, the mixed-step algorithm has the same order of rate of convergence as the original procedure. In particular, the error order is minimized for $\rho = 1/3$ and is proportional to $N^{-\frac{1}{3}}$. In order to go further in the comparison with the previous algorithm, it is now natural to compare the variance $\hat{\sigma}^2_T$ with

$$\hat{\sigma}^2_0 = \int |\sigma^*\nabla g_{_0}(x)|^2\nu(dx) = -2\int g_{_0}(x)\mathcal{A}g_{_0}(x)\nu(dx).$$

which denotes that of the original procedure where $g_{_0}$ is a solution to the Poisson equation: $f - \nu(f) = -\mathcal{A}g_{_0}$. As a first example, we focus on a very particular case: the one-dimensional Ornstein-Uhlenbeck

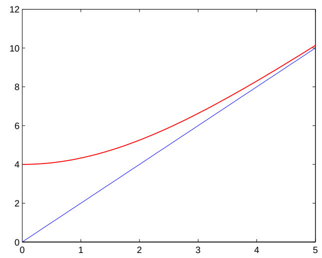

Figure 1: $T \mapsto \hat{\sigma}_T^2$ for the $O.U.$ process with $f(x) = x^2$

process $dX_t = -\frac{1}{2}X_t + dW_t$ with the function $f(x) = x^2$. In this case, closed forms are available, namely

$$\hat{\sigma}_0^2 = 4 \quad \text{and } \forall T > 0, \quad \hat{\sigma}_T^2 = \frac{2T(1+e^{-T})}{1-e^{-T}} \geq 4.$$

$T \to \hat{\sigma}_T^2$ is a continuous increasing function on $\mathbb{R}_+$ with linear growth (see Figure 1).
In a general setting, some of the above properties can be preserved. For instance, under the (nice) assumptions of Section 7, it can be shown that for every $\mathcal{C}^2$-function f such that f, ∇f and $D^2 f$ are bounded, the following properties hold:

$$\lim_{T\to 0} \hat{\sigma}_T^2 = \hat{\sigma}_0^2, \quad \lim_{T\to 0} \partial_T \hat{\sigma}_T^2 = 0 \quad \text{and} \quad \lim_{T\to+\infty} \partial_T \hat{\sigma}_T^2 = \int (f - \nu(f))^2 \nu(dx). \tag{3.15}$$

These properties are proved in Appendix B. Roughly speaking, (3.15) says that the variance is close to that of the original procedure near 0 and that, asymptotically, $T \mapsto \hat{\sigma}_T^2$ increases linearly with T with a gradient being the variance of f under the invariant distribution.

3.3 Functional case

We now consider the problem of the computation of $\mathbb{E}_\nu(F) = \int \mathbb{E}[F(X^x)]\nu(dx)$ where $F : \mathcal{C}([0,T], \mathbb{R}^d) \to \mathbb{R}$ is a Lipschitz continuous functional (with a slight abuse of notation, for $\alpha \in \mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$, we will write $F(\alpha)$ instead of $F(\alpha_t, 0 \leq t \leq T)$). We denote by f_F the function defined for every $x \in \mathbb{R}^d$ by

$$f_F(x) = \mathbb{E}[F(X^x)].$$

Like in the marginal case, the rate of convergence of the procedure is strongly relateded to the weak error over a finite horizon, *i.e.* to the error between $\mathbb{E}[F(X^x)]$ and $\mathbb{E}[F(\xi^{x,\mathbf{h}})]$. However, by contrast with the marginal case, there is no general expansion of such an error in the functional setting. That is why we introduce the following assumption:
$(\mathbf{C_F}(\alpha))$ $(\alpha \in]\frac{1}{2}, 1])$: For any sequence of positive numbers $\mathbf{h} := (h_k)_{k=1}^{k_T}$ such that $\sum_{k=1}^{k_T} h_k = T$,

$$\Big| \mathbb{E}[F(X^x)] - \mathbb{E}[F(\xi^{x,\mathbf{h}})] \Big| \leq C(1 + |x|^r) \|\mathbf{h}\|_\infty^\alpha$$

where $\|\mathbf{h}\|_\infty = \max_{k=1}^{k_T} h_k$ and r is a positive number.

In fact, when F is a Lipschitz continuous functional, the above assumption is true with $\alpha = 1/2$ and $r = 1$ (since $\mathbb{E}[\sup_{0\le t\le T} |X_t^x - \xi_t^{x,\mathbf{h}}|] = O(\sqrt{h})$). Case $\alpha \in (1/2, 1]$ can be of interest for some particular functionals. For instance, if $F(\alpha) = f(\int_0^T \alpha_s ds)$ where f is Lispchitz continuous, it can be shown that $(\mathbf{C_F(1)})$ holds (see [12]). The below theorem is then divided into two parts, respectively without or with this additional assumption:

Theorem 3.2. *Assume* $(\mathbf{S_{a,\infty}})$ *holds with an* $a \in (0,1]$. *Assume* $(\mathbf{S_T^\nu})$. *Let* $\alpha \in (0,1]$ *such that* b *and* σ *are* $\mathcal{C}^{4,\alpha}$*-functions on* $\mathbb{R}^d$ *with bounded derivatives. Let* $F : \mathcal{C}([0,T],\mathbb{R}^d) \to \mathbb{R}$ *be a Lipschitz continuous function such that* f_F *satisfies* $(\mathcal{P}^{\mathrm{pol}}_{\mathbf{4,T}})$ *with a function denoted by* g_F. *Then,*

(i) If $\frac{1}{\sqrt{n}}\sum_{k=1}^n \sqrt{\gamma_{N(kT)}} \xrightarrow{n\to+\infty} 0$,

$$\sqrt{n}\left(\bar{\mu}^{(n)}(F) - \mathbb{E}_\nu(F)\right) = \sqrt{n}\left(\frac{1}{n}\sum_{k=1}^n F(\xi^{((k-1)T)}) - \mathbb{E}_\nu(F)\right) \stackrel{n\to+\infty}{\Longrightarrow} \mathcal{N}\left(0;\widetilde{\sigma}_F^2\right)$$

with $\widetilde{\sigma}_F^2 = \int \left(F(X^x) - \mathbb{E}[F(X^x)] + \frac{1}{T}\left(g_F(X_T^x) - \mathbb{E}[g_F(X_T^x)]\right)\right)^2 \nu(dx)$.

(ii) If $(\mathbf{C_F}(\alpha))$ *holds for an* $\alpha \in]\frac{1}{2},1]$ *and* $\frac{1}{\sqrt{n}}\sum_{k=1}^n \left(\gamma_{N(kT)}\right)^\alpha \xrightarrow{n\to+\infty} 0$, *the above weak rate of convergence of* $\bar{\mu}^{(n)}$ *remains true.*

Remark 3.3. Let $\alpha \in \left[\frac{1}{2},1\right]$ and set $\gamma_n = \gamma_1 n^{-\rho}$. Owing to the computations of Section 3.2, we have

$$\frac{1}{\sqrt{n}}\sum_{k=1}^n (\gamma_{N(kT)})^\alpha \xrightarrow{n\to+\infty} 0 \iff \rho > \frac{1}{2\alpha+1}$$

and the (weak) rate of convergence, written this time in terms of the number $N = N(nT)$ of discretization times of the Euler scheme, is of order

$$\Gamma_N^{-\frac{1}{2}} \sim CN^{-\frac{1-\rho}{2}} \quad \text{if } \rho \in \left(\frac{1}{2\alpha+1}, 1\right).$$

Note in particular, when $\alpha = 1/2$, this leads to an "optimal" error which is $O(N^{\frac{1}{2}+\varepsilon})$ for every $\varepsilon > 0$. Again, the rate order $\Gamma_N^{-\frac{1}{2}}$ is the same as that obtained for the weighted empirical measure $\nu^{(N)}(\xi(\omega), d\alpha)$ originally introduced in [17]. However, for a given functional $F : \mathcal{C}([0,T],\mathbb{R}^d) \to \mathbb{R}$, the computation of $\bar{\mu}^{(n)}(\omega, F)$ is (*a priori* much) less demanding in terms of complexity than that of $\nu^{(N)}(\xi(\omega), d\alpha)$. More precisely, the complexity of $(\bar{\mu}^{(n)}(\omega, F))_{n\ge1}$ is linear in terms of the number N of discretization times, *i.e.* the number of operations for the computation of $\bar{\mu}^{(n)}(\omega, F)$ is proportional to

$$\kappa_E \times N(nT) + \sum_{k=1}^n \kappa_F \times (N(kT) - N((k-1)T) = (\kappa_E + \kappa_F) \times N(nT)$$

where κ_E denotes the computational complexity of one iteration of the Euler scheme and $\kappa_F \times m$ the complexity of the computation of $F(\xi)$ where $\xi \in \mathcal{C}([0,T],\mathbb{R}^d)$ is a typical path of an Euler scheme with m time steps. For the original procedure a similar computation leads to

$$\kappa_E \times N(nT) + \sum_{k=0}^{N(nT)-1} \kappa_F(N(\Gamma_k + T) - N(\Gamma_k)) \approx \kappa_E \times N(nT) + \kappa_F \sum_{\ell=1}^n (N(\ell T) - N((\ell-1)T))^2$$

which grows to infinity infinitely faster. However in some particular cases (fortunately often those of interest) the functional F can be computed recursively (see [18] for details) so that our estimate $\kappa_F \times m$ should be replaced by $\kappa_{_F} \times c_0$ which makes the resulting global complexity comparable to our new approach. As a conclusion, our new procedure yields a better control of the complexity for all (computable) functionals F.

In practice, $F(\xi^{(kT)})$ is not always computable since the genuine Euler scheme is non-constant on any time interval. If we replace the genuine Euler scheme by its stepwise constant càdlàg counterpart, it can be shown that the CLT established in (i) still holds for the same functional F under the slightly more stringent assumption that for every $\varepsilon > 0$, $\frac{1}{\sqrt{n}}\sum_{k=1}^{n}(\gamma_{N(kT)})^{\frac{1}{2}-\varepsilon} \xrightarrow{n\to+\infty} 0$.

3.4 The Richardson-Romberg extrapolation

In this section, we want to perform and analyze a Richardson-Romberg (RR) extrapolation to cancel the first order term in the expansion of the time discretization. Note that we will only consider this problem in the marginal case since we do not have any explicit expansion of the weak error in the functional case. As concerns discretization schemes with decreasing step this procedure has been introduced in [14] (see Chapter V). However, our proposal is slightly different since we will design our two schemes in a more consistent way.

Let $\theta \in [0,1]$ and let $(W, \widetilde{W})$ denote a $2q$-dimensional SBM. Denote again by $(\mathcal{F}_t)_{t\geq 0}$ the usual augmentation of $(\sigma((W,\widetilde{W})_s, 0 \leq s \leq t))_{t\geq 0}$. Then, we define W^θ by

$$W^\theta = \theta W + \sqrt{1-\theta^2}\widetilde{W} \quad \text{so that} \quad \langle W, W^\theta\rangle_t = \theta t.$$

We denote by $(\xi_t^{(\theta)})_{t\geq 0}$ the Euler scheme built with the increments of W^θ and a step sequence $(\widetilde{\gamma}_n)_{n\geq 1}$ defined for every $n \geq 1$ by

$$\widetilde{\gamma}_{2n-1} = \widetilde{\gamma}_{2n} = \frac{\gamma_n}{2}.$$

In the sequel, we set $\widetilde{\Gamma}_n = \sum_{k=1}^{n}\widetilde{\gamma}_k$ (note that $\widetilde{\Gamma}_{2n} = \Gamma_n$) and for every $t \geq 0$, $\widetilde{N}(t) = \inf\{k \geq 0, \widetilde{\Gamma}_{k+1} > t\}$.

We also denote by $(\bar{\mu}_n^{(\theta)}(\omega, dy))_{n\geq 1}$, the sequence of empirical measures related to $(\xi_t^{(\theta)})_t$ defined for every $n \geq 1$ by

$$\bar{\mu}_n^{(\theta)}(\omega, f) = \frac{1}{n}\sum_{k=1}^{n} f(\xi_{(k-1)T}^{(\theta)}(\omega)).$$

Finally, we consider the sequence of (weighted) random measures $(\widetilde{\mu}_n^{(\theta)}(\omega, dy))_{n\geq 1}$ defined by the extrapolation of $\bar{\mu}_n$ and $\bar{\mu}_n^{(\theta)}$ in proportion with scaling factor of their respective step sequences:

$$\widetilde{\mu}_n^{(\theta)}(\omega, f) = 2\bar{\mu}_n^{(\theta)}(\omega, f) - \bar{\mu}_n(\omega, f).$$

In the sequel of this section, we are then interested in the rate of convergence of $(\widetilde{\mu}_n^{(\theta)}(\omega, f))_{n\geq 1}$ toward $\nu(f)$. To this aim , we introduce $\mathbb{X}^{(\theta)} := (X, X^{(\theta)})$ solution to the system of *duplicated* SDE's

$$\begin{cases} dX_t = b(X_t)dt + \sigma(X_t)dW_t \\ dX_t^{(\theta)} = b(X_t^{(\theta)})dt + \sigma(X_t^{(\theta)})dW_t^\theta \end{cases} \tag{3.16}$$

starting from $(X_0, X_0^{(\theta)})$ $\mathbb{R}^{2d}$-valued random vector supposed to be independent of $\sigma(W, \widetilde{W})$). Since b and σ are Lipschitz continuous functions, the $\mathbb{R}^{2d}$-valued process $\mathbb{X}^{(\theta)}$ is a Feller Markov diffusion process whose semi-group is denoted by $(Q_t^{(\theta)})_{t\geq 0}$. In the theorem below, we need that $Q_T^{(\theta)}$ has a *unique* invariant distribution $\nu^{(\theta)}$. It is clear that, if P_T has a unique invariant distribution ν, then $\nu^{(\theta)}$ belongs to $\mathcal{C}_{\nu,\nu} := \{\mu \in \mathcal{P}(\mathbb{R}^{2d}), \mu_1 = \mu_2 = \nu\}$ where $\mathcal{P}(\mathbb{R}^{2d})$ denotes the set of probabilities on $\mathbb{R}^{2d}$ and $\mu_1 = \mu(\,.\,\times \mathbb{R}^d)$ and $\mu_2 = \mu(\mathbb{R}^d \times\,.\,)$ (first and second marginals of μ). Since $\mathcal{C}_{\nu,\nu}$ is a weakly compact and convex subset of the Banach space of signed measures on $(\mathbb{R}^{2d}, \mathcal{B}(\mathbb{R}^{2d}))$, the existence of an invariant distribution $\nu^{(\theta)}$ for $Q_T^{(\theta)}$ follows from the Kakutani fixed-point Theorem applied to $\mu \mapsto \mu Q_T^{(\theta)}$ (in fact dealing with the temporal mean of $(Q_t^{(\theta)})_{t\geq 0}$ would yield an invariant distribution for the whole semi-group if needed). Moreover,we know that uniqueness of $\nu^{(\theta)}$ implies that of ν still owing to the Kakutani Theorem since, if ν and $\tilde{\nu}$ were two invariant distributions of P_T then $\mu \mapsto \mu Q_T^{(\theta)}$ would have another fixed distribution on $\mathcal{C}_{\nu,\tilde{\nu}}$. The reverse is not true but uniqueness of $\nu^{(\theta)}$ being important to this problem, it will be discussed in Section 3.5 below.

Let us now state the main result of this section:

Theorem 3.3. *Assume* $(\mathbf{S}_{\mathbf{a},\infty})$ *holds with an* $a \in (0,1]$ *and assume that* $\mathrm{Tr}(\sigma\sigma^*(x)) = o(V^a(x))$ *as* $|x| \to +\infty$. *Let* $\theta \in [0,1]$ *and assume that* $Q_T^{(\theta)}$ *admits a unique invariant distribution* $\nu^{(\theta)}$. *Assume that* b *and* σ *are* $\mathcal{C}^{9,\alpha}$*-functions on* $\mathbb{R}^d$ *with bounded existing derivatives,* $\alpha \in (0,1]$. *Let* $f : \mathbb{R}^d \to \mathbb{R}$ *be a Borel function satisfying* $(\mathcal{P}^{\mathrm{pol}}_{\mathbf{9},\mathbf{T}})$ *and denote by* g_T *the solution to* (3.11). *Then, assume that* $\varphi^{(1)}$ *defined by* $\varphi^{(1)}(x) = \frac{1}{2}\int_0^T \mathbb{E}[\Phi_{g_T}(X_s^x)]ds$ *satisfies* $(\mathcal{P}^{\mathrm{pol}}_{\mathbf{5},\mathbf{T}})$ *and denote by* $g_{\varphi^{(1)}}$ *the associated solution to* (3.11). *Assume that* $(\gamma_n)_{n\geq 1}$ *satisfies* $\sum_{k=1}^{+\infty} \gamma^2_{N(kT)} = +\infty$ *and* $\gamma_{N(nT)} - \gamma_{N((n+1)T)} = o(\gamma^2_{N(nT)})$. *Then,*

(i) If $\frac{1}{\sqrt{nT}} \sum_{k=1}^n \gamma^2_{N(kT)} \xrightarrow{n\to+\infty} 0$, *then*

$$\sqrt{nT}\left(\widetilde{\mu}_n^{(\theta)}(\omega, f) - \nu(f)\right) \stackrel{n\to+\infty}{\Longrightarrow} \mathcal{N}\Big(0, (\hat{\sigma}_T^{(\theta)})^2\Big)$$

$$\text{where,}\quad (\hat{\sigma}_T^{(\theta)})^2 = 5\hat{\sigma}_T^2 - \frac{4}{T}\int \Big(g_T(x)g_T(y) - P_Tg_T(x)P_Tg_T(y)\Big)\nu^{(\theta)}(dx,dy). \tag{3.17}$$

(ii) If $\sqrt{\frac{T}{n}} \sum_{k=1}^n \gamma^2_{N(kT)} \xrightarrow{n\to+\infty} \beta_0 \in (0,+\infty]$, *then* $\frac{n}{\sum_{k=1}^n \gamma^2_{N(kT)}}\Big(\widetilde{\mu}_n^{(\theta)}(\omega, f) - \nu(f)\Big)$ *is tight with bounded (resp. subgaussian) weak limits if* $\beta_0 = +\infty$ *(resp.* $\beta_0 < +\infty$*).*

Furthermore, if $(\gamma_n)_{n\geq 1}$ *is such that* $N(nT) = nT$ *for every* $n \geq 1$, *then we have the following more precise result:*

– If $\beta_0 = +\infty$,

$$\frac{n}{\sum_{k=1}^n \gamma^2_{N(kT)}}\left(\widetilde{\mu}_n^{(\theta)}(\omega, f) - \nu(f)\right) \xrightarrow{\mathbb{P}} \widetilde{m}_T = -\frac{1}{2T}\int_{\mathbb{R}^d}\int_0^T \mathbb{E}\left[\frac{1}{T}\Phi_{g_{\varphi^{(1)}}}(X_s^x) + \chi_g(s, X_s^x)\right] ds\nu(dx)$$

as $n \to +\infty$, *where* χ_g *is given by* (4.30).

– If $\beta_0 \in (0, +\infty)$,

$$\sqrt{nT}\left(\widetilde{\mu}_n^{(\theta)}(\omega, f) - \nu(f)\right) \stackrel{n\to+\infty}{\Longrightarrow} \mathcal{N}\left(\beta_0\,\widetilde{m}_T, (\hat{\sigma}_T^{(\theta)})^2\right).$$

Remark 3.4. • Using the stationarity property, $(\hat{\sigma}_T^{(\theta)})^2$ can be written as follows:

$$(\hat{\sigma}_T^{(\theta)})^2 = \frac{1}{T}\int_{\mathbb{R}^d\times\mathbb{R}^d} \mathrm{Var}_{x,y}\left(2g_T(X_T^{(\theta)}) - g_T(X_T)\right)\nu^{(\theta)}(dx,dy).$$

• Owing to the definition of $N(t)$, we have for every $k\in\mathbb{N}$,

$$C_T^1\gamma^2_{N(k+1)T} \leq \sum_{l=N(kT)+1}^{\gamma_{N((k+1)T)}} \gamma_l^3 \leq C_T^2\gamma^2_{N(kT)}, \tag{3.18}$$

so that

$$\sum_{k=1}^{n}\gamma^2_{N(kT)} \xrightarrow{n\to+\infty} +\infty \iff \sum_{k\geq 1}\gamma_k^3 = +\infty.$$

If $\gamma_n = \gamma_1 n^{-\rho}$ with $\rho\in(0,1)$, this condition is satisfied if and only if $\rho < 1/3$. Using the computations of Subsection 3.2, one checks that in this case, $\gamma_{N(nT)} - \gamma_{N((n+1)T)} = o(\gamma^2_{N(nT)})$, and that,

$$\frac{1}{\sqrt{n}}\sum_{k=1}^{n}\gamma^2_{N(kT)} \xrightarrow{n\to+\infty} 0 \iff \rho > \frac{1}{4}.$$

In fact, the Romberg extrapolation implies that the discretization error is lower than in Theorem 3.1 and thus, that the rate of convergence in (i) can be preserved for smaller values of ρ. More precisely, after N discretization times, the error is of order

$$\begin{cases} N^{-\frac{1-\rho}{2}} & \text{if } \rho\in(1/5,1/3) \\ N^{-\rho} & \text{if } \rho\in(0,1/5]. \end{cases}$$

The order of the error attains its minimum at $\rho = 1/5$ and is then proportional to $N^{-\frac{2}{5}}$.

• If the additional assumption on the uniqueness of the invariant distribution for $Q_T^{(\theta)}$ fails, tightness results are preserved. In particular, in (i), we would have only obtained that $\sqrt{n}\left(\widetilde{\mu}_n^{(\theta)}(\omega,f) - \nu(f)\right)$ is a tight sequence with Gaussian limits.

• The last part of the above theorem requires an additional assumption on the steps: $N(nT) = nT$. In fact, if this assumption fails, some edge terms must be managed, especially in the expansion of the weak error relative to this problem, which become non-negligible at the second order. However, it seems that this assumption could be avoided owing to a sharper and tedious study of these edge terms. This refinement will not be tackled in the paper.

An important numerical question relative to Theorem 3.3 is the variance $(\hat{\sigma}_g^{(\theta)})^2$. When one performs a Romberg extrapolation, it is important to compare the variance of the modified scheme with that of the original one. In [16], this problem is tackled for the standard Euler scheme on a finite interval and it is shown that the variance of the original scheme can be preserved by taking $\theta = 1$, *i.e.* by considering the same underlying Brownian motion for both schemes. Here, we obtain a similar result for the invariant distribution up to a uniqueness assumption for the invariant distribution when $\theta = 1$. In that case, the distribution ν_Δ defined by $\nu_\Delta(h) = \int h(x,x)\nu(dx)$ is clearly an invariant distribution for $Q_T^{(0)}$. If this is the only one, we can minimize the variance. This is the purpose of the next proposition.

Proposition 3.3. *Assume* $(\mathbf{S}^{\nu}_{\mathbf{T}})$. *Let* $f \in L^2(\nu)$.
(i) For every $\theta \in [0,1]$, $\hat\sigma^2_T \le (\hat\sigma^{(\theta)}_T)^2$.
(ii) If ν_Δ *is the unique invariant distribution of* $Q^{(1)}_T$, *then* $(\hat\sigma^{(1)}_T)^2 = \hat\sigma^2_T$.

Proof. (i) Owing to the stationarity property, we have

$$\begin{aligned}\int g(x)g(y) - &P_T g(x) P_T g(y) \nu^{(\theta)}(dx, dy)\\ &= \int \mathbb{E}_{x,y}\left[\Big(g(X_T) - \mathbb{E}[g(X_T)]\Big)\Big(g(X^{(\theta)}_T) - \mathbb{E}[g(X^{(\theta)}_T)]\Big)\right]\nu^{(\theta)}(dx, dy)\\ &= \mathbb{E}_{\nu^{(\theta)}}\left[\Big(g(X_T) - P_T g(X_0)\Big)\Big(g(X^{(\theta)}_T) - P_T g(X^{(\theta)}_0)\Big)\right].\end{aligned}$$

Using that under the stationary regime, $\mathcal{L}((X_t)_{t\ge 0}) = \mathcal{L}((X^{(\theta)}_t)_{t\ge 0}) = \mathbb{P}_\nu$, we derive from Schwarz's inequality that

$$\int \Big(g(x)g(y) - P_T g(x) P_T g(y)\Big)\nu^{(\theta)}(dx, dy) \le \mathbb{E}_\nu\left[(g(X_T) - P_T g(X_0))^2\right] = \hat\sigma^2_T \tag{3.19}$$

Owing to (3.17), this shows that for every $\theta \in [0,1]$, $(\hat\sigma^{(\theta)}_T)^2 \ge \hat\sigma^2_T$.
(ii) Set

$$U = g(X_T) - P_T g(X_0) \quad \text{and} \quad V = g(X^{(\theta)}_T) - P_T g(X^{(\theta)}_0).$$

From what preceeds and from the equality case in Schwarz's inequality, we obtain that equality holds in (3.19) if and only if there exists $\lambda \in \mathbb{R}$ such that $U = \lambda V$ *a.s.*, and thus if and only if $U = V$ *a.s.* since U and V have the same distribution. As a consequence, equality holds in (3.19) if $(X_0, X_T) = (X^{(\theta)}_0, X^{(\theta)}_T)$ *a.s.* So is the case if $\nu^{(\theta)} = \nu_\Delta$. □

3.5 Are invariant measures of duplicated diffusions always supported by the diagonal?

This section is a summarized version of [15]. In particular, we refer to this paper for the proofs of the results.
We keep the notations of the previous part of the paper: denoting by $(X_t)_{t\ge 0}$ the solution to (1.1) and by $(\mathbb{X}^{(\theta)}_t)_{t\ge 0}$ the *duplicated* diffusion solution to (3.16), we consider the problems of existence and mostly of uniqueness of the invariant distribution for $(\mathbb{X}^{(\theta)}_t)_{t\ge 0}$.
In what follows we will always assume that the original diffusion X^x has at least one invariant distribution denoted ν *i.e.* satisfying $\nu P_t = \nu$ for every $t \in \mathbb{R}_+$.

▷ *Existence of an invariant distribution for* $(Q^{(\theta)}_t)_{t\ge 0}$ *with marginals* ν. Using that the family of probability measures on $(\mathbb{R}^d \times \mathbb{R}^d, \mathcal{B}or(\mathbb{R}^d)^{\otimes 2})$

$$\frac{1}{t}\int_0^t \nu^{\otimes}(dx_0, dx'_0) Q^{(\theta)}_s(x_0, x'_0, dy, dy')ds, \; t > 0$$

is tight since both its marginals on $\mathbb{R}^d$ are equal to ν and that the semi-group $(Q^{(\theta)}_t)_{t\ge 0}$ is Feller, one easily shows that any limiting distribution $\nu^{(\theta)}$ as $t \to \infty$ is an invariant distribution for $(Q^{(\theta)}_t)_{t\ge 0}$ with both marginals equal to ν.

$\triangleright$ *Uniqueness of the invariant distribution of $Q_T^{(\theta)}$ (and thus for $(Q_t^{(\theta)})_{t\geq 0}$).* It is clear that in full generality $(\mathbb{X}_t^{(\theta)})_{t\geq 0}$ may admit several invariant distributions even if X has only one such distribution. So is the case in a fully degenerate setting: $\sigma \equiv 0$. Then, if the flow $\Phi(x,t)$ of the ODE $\dot{x} = b(x)$ has an invariant distribution ν, then both $\nu^{\otimes 2}$ and the image ν_Δ of ν on the diagonal of $\mathbb{R}^d \times \mathbb{R}^d$ are invariant for the duplicated o.d.e. and $\nu^{\otimes 2}$ and the image ν_Δ on the diagonal of $\mathbb{R}^d \times \mathbb{R}^d$ are invariant and $\nu^{\otimes 2} \neq \nu_\Delta$ as soon as ν is not reduced to Dirac mass (think for instance to a 2-dimensional ODE with a limit cycle).

In the non-degenerate case ($\sigma \not\equiv 0$), the situation is more involved and depends on the correlation θ between the two Brownian motions.

• $\theta \in [0,1)$: if one writes $W^{(\theta)} = \theta W + (1-\theta^2)^{\frac{1}{2}}\widetilde{W}$ where $\widetilde{W}$ is independent of W, one checks that for every $t \geq 0$, the diffusion matrix, $\Sigma_{\mathbb{X}_t^{(\theta)}}$ of the duplicated diffusion $\mathbb{X}^{(\theta)} = (X, X^{(\theta)})$ at time t satisfies

$$\Sigma_{\mathbb{X}_t^{(\theta)}}\Sigma^*_{\mathbb{X}_t^{(\theta)}} = \left[\begin{array}{cc} \sigma\sigma^*(X_t) & \theta\sigma(X_t)\sigma^*(X_t^{(\theta)}) \\ \theta\sigma(X_t)\sigma^*(X_t^{(\theta)}) & \sqrt{1-\theta^2}\sigma\sigma^*(X_t^{(\theta)}) \end{array}\right].$$

From this expression, it is straightforward than ellipticity or uniform ellipticity of σ (for X^x which underlines $q \geq d$) can be transferred to the couple $(X, X^{(\theta)})$. Uniform ellipticity classically implies under standard regularity and growth/boundedness assumptions on the coefficients b, σ and their partial derivatives the existence of a (strictly) positive probability density $p_t(x,y)$ for X. These conditions once again transfer to the coefficients of the diffusion $(X, X^{(\theta)})$. Consequently, under these standard assumptions on b and σ which ensure uniqueness of the invariant distribution ν for X (and more precisely for P_T), we get uniqueness for the "duplicated" diffusion process $(X, X^{(\theta)})$ (and more precisely for $Q_T^{(\theta)}$) as well.
In an hypoelliptic setting, it is also classical that if the diffusion coefficient of a diffusion satisfies the hypoelliptic Hörmander assumption and if the deterministic system related to the Stochastic differential system (written in the Stratonovich sense) is controllable, then (the transition semi-group is strongly Feller and) the invariant distribution, if any, is unique. In fact, both properties can be transferred from the original SDE to the duplicated system so that the invariant distribution is unique (see [15] for details).

• $\theta = 1$: this setting corresponds in our problem to the simulation of consistent Brownian increments for our Euler schemes. Owing to Proposition 3.3, in this case, uniqueness of the invariant distribution for $Q_T^{(1)}$ implies that the variance in the CLT is equal to that of the original procedure. However, the situation becomes significantly different since the resulting "duplicated stochastic system" becomes degenerate. It is clear that ν_Δ as defined above is an invariant distribution for the system, the question becomes: "is it the only one?".

In what follows, we provide some answers: In Proposition 3.4, we state that under some asymptotic confluence assumptions, uniqueness holds for the invariant distribution of the "single" diffusion and is transferred to the duplicated diffusion. Then, when this assumption fails and if $d \geq 2$, we show explicitely that it is possible to build some diffusions for which uniqueness holds for the "single" but not for the duplicated one.
Finally, we focus on the one-dimensional case where the property is essentially always true.

Proposition 3.4 (see [15]). *Assume that the SDE is* asymptotically confluent *in the following* *sense:*

$$(AC)\ \equiv\ \forall\, R>0,\quad \exists \delta_R>0,\quad \textit{such that}\ \forall\, x,y\in B_{|.|}(0;R),\ 0<|x-y|\le\delta_R \Longrightarrow$$
$$\langle b(x)-b(y),x-y\rangle+\frac{1}{2}\mathrm{Tr}\Big((\sigma(x)-\sigma(y)(\sigma(x)-\sigma(y)^*\Big)<0,$$

then ν and ν_Δ are respectively the unique invariant distributions for $P_{_T}$ and $Q_T^{(\theta)}$.

Remark 3.5. • In [15], the reader will find other asymptotic confluence conditions which ensure the conclusion of Proposition 3.4.

• If b and σ are differentiable, Assumption (AC) is equivalent to

$$\forall\, R>0,\ \exists\, \varepsilon_R>0 \text{ such that }\ , |x|\le R\Longrightarrow J_b(x)+\frac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{q}\Big(\sum_{k=1}^{d}\frac{\partial\sigma_{ij}}{\partial x_k}(x)\Big)^2<-\varepsilon_R$$

where J_b denotes the Jacobian $d\times d$ matrix of b.

As soon as $d\ge 2$, there are examples of non asymptotically confluent diffusions having a unique invariant distribution ν but whose "duplicated system" (with the same Brownian motion) has several invariant distributions. One example is provided below.

A counterexample: We consider the 2-dimensional SDE with Lipschitz continuous coefficients defined by

$$\begin{aligned} b(x) &= \Big(x\mathbf{1}_{\{0\le|x|\le 1\}}-\frac{x}{|x|}\mathbf{1}_{\{|x|\ge 1\}}\Big)(1-|x|)\\ \sigma(x) &= \vartheta \mathrm{Diag}(b(x))+\begin{bmatrix} 0 & -cx^2\\ cx^1 & 0\end{bmatrix}. \end{aligned}$$

where ϑ, $c\in(0,+\infty)$ are fixed parameters. Switching to polar coordinates $X_t=(r_t\cos\varphi_t, r_t\sin\varphi_t)$, $t\in\mathbb{R}_+$ show that this SDE also reads

$$\begin{aligned} dr_t &= \min(r_t,1)(1-r_t)(dt+\vartheta dW_t^1),\quad r_0\in\mathbb{R}_+\\ d\varphi_t &= c dW_t^2,\quad \varphi_0\in[0,2\pi). \end{aligned}$$

where $x_0=r_0(\cos\varphi_0,\sin\varphi_0)$ and $W=(W^1,W^2)$ is a standard 2-dimensional Brownian motion.

Standard considerations about Feller classification (see [7], chapter 15.6, p. 226) show that, if $x_0\neq 0$ (*i.e.* $r_0>0$) and $\vartheta\in(0,\sqrt{2})$ then

$$r_t\longrightarrow 1\ \text{ as } t\to\infty.$$

while it is classical background that

$$\mathbb{P}\text{a.s.}\quad \forall\varphi_0\in\mathbb{R}_+,\quad \frac{1}{t}\int_0^t\delta_{e^{i(\varphi_0+cW_s^2)}}ds\Longrightarrow\lambda_{S^1}\qquad\text{as } t\to\infty$$

where S^1 denotes the unit circle of $\mathbb{R}^2$ and λ_{S_1} denotes the Lebesgue measure on S^1. Combining these two results straightforwardly yields

$$\forall\, x\in\mathbb{R}^2\setminus\{(0,0)\},\ \mathbb{P}\text{-}a.s.\quad \frac{1}{t}\int_0^t\delta_{X_s^x}ds\overset{(\mathbb{R}^2)}{\Longrightarrow}\nu:=\lambda_{S^1}\qquad\text{as}\qquad t\to\infty.$$

On the other hand, given the form of φ_t, it is clear that if $x = r_0 e^{i\varphi_0}$ and $x' = r'_0 e^{i\varphi'_0}$, r_0, $r'_0 \neq 0$ then

$$\lim_{t\to\infty} |X_t^x - X_t^{x'}| = |e^{i(\varphi_0-\varphi'_0)} - 1|.$$

Now, taking $\varphi_0 \neq \varphi'_0$ and using that the weak limits of $(\frac{1}{t}\int_0^t Q_s(x,x',.,.)ds)_{t\geq 0}$ always belong to the set of invariant distributions of $(Q_t)_{t\geq 0}$, one easily deduces that ν_Δ (defined by $\nu_\Delta(h) = \int h(x,x)\nu(dx)$) can not be the only invariant distribution : otherwise, setting $f_M(x,y) = |y-x| \wedge M$ $(M>0)$, one would have $\frac{1}{t}\int_0^t Q_s f_M(x,x')ds \to 0$ for every $M>0$).

In fact, one can also directly check (with the definition of the semi-group) that $\nu\otimes\nu$ is another invariant distribution for $(Q_t)_{t\geq 0}$ (so that ν_Δ, $\nu\otimes\nu$ and all the convex combinations of these probability measures are invariant distributions for $(Q_t)_{t\geq 0}$).

$\triangleright$ The one-dimensional case. Finally, one has a more unexpected result in one dimension $d = q = 1$.

Theorem 3.4 (see [15]). *Assume that b and σ are locally Lipschitz continuous on $\mathbb{R}$ and that there exists a $\mathcal{C}^2$-function $V : \mathbb{R}^d \to \mathbb{R}_+^*$ such that $\lim_{|x|\to+\infty} V(x) = +\infty$ and $\mathcal{A}V \leq CV$ with $C>0$. Assume that $(P_t)_{t\geq 0}$ admits a unique invariant distribution. Then, $\nu_\Delta = \nu \circ (\xi \mapsto (\xi,\xi))^{-1}$ is the unique invariant distribution of $(Q_t^{(1)})_{t\geq 0}$.*

Remark 3.6. When σ never vanishes, this result can be retrieved by introducing the scale function p of the diffusion defined by:

$$p(x) = \int_{x_0}^x d\xi e^{-\int_{x_0}^{\xi} \frac{2b}{\sigma^2}(u)},\ x \in \mathbb{R}.$$

In fact, by a result of Has'minskii (see [6], Corollary A.2 p.274), when the diffusion is positive recurrent (see [15] for more detailed assumptions), then, for every $x,y \in \mathbb{R}^d$, $p(X_t^x) - p(X_t^y) \xrightarrow{t\to+\infty} 0$ *a.s.* Denoting by μ an invariant distribution of $(X_t^x, X_t^y)_{t\geq 0}$, this implies that

$$\forall K \geq 0, \quad \int |p(x)-p(y)| \wedge K\mu(dx,dy) \leq \limsup \frac{1}{t}\int_0^t \mathbb{E}_\mu[|p(X_s^x) - p(X_s^y)| \wedge K]ds = 0$$

As a consequence, $p(x) = p(y)$ $\mu(dx,dy)$-*a.s.* Since p is an increasing function, we can deduce that $\mu((x,x), x\in\mathbb{R}^d) = 1$ and thus that $\mu = \nu_\Delta$.

The next three sections are devoted to the proofs of the main results (of Section 3). In Section 4, we begin with some first and second order expansions of the weak error related to the Euler scheme in a finite horizon in a non-constant step framework. These expansions are important tools for the sequel of the proof.

4 The weak error in a non constant-step framework

In the sequel, C denotes any non explicit positive constant. Likewise, for a function f with polynomial growth, we usually write $|f(x)| \leq C(1+|x|^r)$ where the exponent r denotes a non explicit positive number.

Let $T>0$. We denote by $(\xi_t^{x,\mathbf{h}})_{t\in[0,T]}$ the genuine Euler scheme with step sequence $\mathbf{h} := (h_k)_{k\geq 1}$ starting point $x\in\mathbb{R}^d$ and by $\mathcal{E}_T(g,x,\mathbf{h})$ the weak error between the Euler scheme and the diffusion, that is:

$$\mathcal{E}_T(g,x,\mathbf{h}) = \mathbb{E}[g(\xi_T^{x,\mathbf{h}})] - \mathbb{E}[g(X_T^x)]. \tag{4.20}$$

4.1 Preliminary lemmas

We denote by (t_k) the sequence of discretization times:

$$t_0 = 0, \quad t_k = \sum_{l=1}^{k} h_l, \quad k \geq 1,$$

and we assume that there exists $k_{_T} \in \mathbb{N}$ such that $t_{k_{_T}} = T$.

Lemma 4.1. *Let $k \in \mathbb{N}$ and let $g : \mathbb{R}^d \to \mathbb{R}$ be such that the assumptions of Proposition 2.2 hold. Then,*
(i) If $m = 4$,

$$|\mathcal{E}_{_T}(g, x, \mathbf{h})| \leq C_T(1 + |x|^r) \sum_{k=1}^{k_{_T}} (h_k)^2$$

where r is a positive real number and C_T is a real constant only depending on T, b, σ, g (and their derivatives up to order 4) which is non-decreasing as a function of T.

(ii) If $m = 5$ or $= 6$, then,

$$\mathcal{E}_{_T}(g, x, \mathbf{h}) = \frac{1}{2} \sum_{k=1}^{k_{_T}} (h_k)^2 \mathbb{E}[\Phi_g(t_{k-1}, \xi^{x,\mathbf{h}}_{t_{k-1}})] + R_T(x, \mathbf{h})$$

where Φ_g is defined by (3.12) and

$$|R_T(x, \mathbf{h})| \leq C_T(1 + |x|^r) \begin{cases} \|\mathbf{h}\|_\infty^{\frac{3}{2}} & \text{if } m = 5 \\ \|\mathbf{h}\|_\infty^2 & \text{if } m = 6 \end{cases}$$

with $r > 0$, $\|\mathbf{h}\|_\infty = \max_{k=1}^{k_{_T}} h_k$, and $T \mapsto C_T$ is a non-decreasing function only depending on T, b, σ, g and their partial derivatives up to order k.

Proof. (i) First, in order to alleviate the notations, we will write u instead of u_g. Then, by the definition of u, we have

$$\mathbb{E}[g(\xi^{x,\mathbf{h}}_T)] - \mathbb{E}[g(X^x_T)] = \sum_{k=1}^{k_T} \mathbb{E}[u(t_k, \xi^{x,\mathbf{h}}_{t_k}) - u(t_{k-1}, \xi^{x,\mathbf{h}}_{t_{k-1}})]. \tag{4.21}$$

By Proposition 2.2, $t \mapsto u(t, x)$ is a $\mathcal{C}^2$-function and $x \mapsto u(t, x)$ is a $\mathcal{C}^4$ function. Then, owing to the Itô formula, we have

$$u(t_k, \xi^{x,\mathbf{h}}_{t_k}) - u(t_{k-1}, \xi^{x,\mathbf{h}}_{t_{k-1}}) = \int_{t_{k-1}}^{t_k} \Big(\partial_t u(s, \xi^{x,\mathbf{h}}_s) + \bar{A}u(s, \xi^{x,\mathbf{h}}_s, \xi^{x,\mathbf{h}}_{\underline{s}})\Big) ds + M^{(0)}_{t_k} - M^{(0)}_{t_{k-1}}$$

where $(M^{(0)}_t)$ is a true martingale defined by $M^{(0)}_t = \int_0^t \langle \partial_x u(s, \xi^{x,\mathbf{h}}_s), \sigma(\xi^{x,\mathbf{h}}_{\underline{s}}) dW_s \rangle$ and

$$\bar{A}u(s, x, \underline{x}) = \langle \partial_x u(s, x), b(\underline{x}) \rangle + \frac{1}{2} \mathrm{Tr}\Big(\sigma^*(\underline{x}) \partial^2_{xx} u(s, x) \sigma(\underline{x})\Big).$$

As a consequence,

$$\mathbb{E}_{t_{k-1}}[u(t_k,\xi^{x,\mathbf{h}}_{t_k})-u(t_{k-1},\xi^{x,\mathbf{h}}_{t_{k-1}})]=\int_{t_{k-1}}^{t_k}\mathbb{E}_{t_{k-1}}[\partial_t u(s,\xi^{x,\mathbf{h}}_s)+\bar{\mathcal{A}}u(s,\xi^{x,\mathbf{h}}_s,\xi^{x,\mathbf{h}}_{\underline{s}})]ds. \tag{4.22}$$

Applying again the Itô formula to $(\partial_t u(t,\xi^{x,\mathbf{h}}_t))_{t\geq 0}$, we have for every $t\in[t_{k-1},t_k]$:

$$\partial_t u(t,\xi^{x,\mathbf{h}}_t)=\partial_t u(t_{k-1},\xi^{x,\mathbf{h}}_{t_k})+\int_{t_{k-1}}^{t}\Big(\partial^2_{tt}u(s,\xi^{x,\mathbf{h}}_s)+\bar{\mathcal{A}}(\partial_t u)(s,\xi^{x,\mathbf{h}}_s,\xi^{x,\mathbf{h}}_{\underline{s}})\Big)ds+M^{(1)}_t-M^{(1)}_{t_{k-1}}$$

where $(M^{(1)}_t)$ is a true martingale defined by $M^{(1)}_t=\int_0^t\langle\partial^2_{xt}u((s,\xi^{x,\mathbf{h}}_s),\sigma(\xi^{x,\mathbf{h}}_{\underline{s}})dW_s\rangle$. Note that

$$\bar{\mathcal{A}}(\partial_t u)(s,x,\underline{x})=\langle\partial^2_{xt}u(s,x),b(\underline{x})\rangle+\frac{1}{2}\mathrm{Tr}\Big(\sigma^*(\underline{x})\partial^3_{xxt}u(s,x).\sigma(\underline{x})\Big).$$

Likewise, we need to develop $\bar{\mathcal{A}}u(t,\xi^{x,\mathbf{h}}_t,\xi^{x,\mathbf{h}}_{\underline{t}})$. We have,

$$\partial_x u(t,\xi^{x,\mathbf{h}}_t)=\partial_x u(t_{k-1},\xi^{x,\mathbf{h}}_{t_{k-1}})+\int_{t_{k-1}}^{t}\Big[(\partial_t+\bar{\mathcal{A}})(\partial_x u)(s,\xi^{x,\mathbf{h}}_s,\xi^{x,\mathbf{h}}_{\underline{s}})\Big]\,ds+M^{(2)}_t-M^{(2)}_{t_{k-1}}$$

and

$$\partial^2_{x^2}u(t,\xi^{x,\mathbf{h}}_t)=\partial^2_{x^2}u(t_{k-1},\xi^{x,\mathbf{h}}_{t_{k-1}})+\int_{t_{k-1}}^{t}(\partial_t+\bar{\mathcal{A}})(\partial^2_{x^2}u)(s,\xi^{x,\mathbf{h}}_s,\xi^{x,\mathbf{h}}_{\underline{s}})ds+M^{(3)}_t-M^{(3)}_{t_{k-1}}$$

where $(M^{(2)}_t)$ and $(M^{(3)}_t)$ are true martingales. Then, plugging the above expansions into (4.22) and using that $(\partial_t u+\mathcal{A}u)(t_{k-1},\xi^{x,\mathbf{h}}_{t_{k-1}})=0$, we obtain

$$\mathbb{E}_{t_{k-1}}[u(t_k,\xi^{x,\mathbf{h}}_{t_k})-u(t_{k-1},\xi^{x,\mathbf{h}}_{t_{k-1}})]=\int_{t_{k-1}}^{t_k}\int_{t_{k-1}}^{s}\mathbb{E}_{t_{k-1}}\bar{\Phi}(v,\xi^{x,\mathbf{h}}_v,\xi^{x,\mathbf{h}}_{\underline{v}})dvds \tag{4.23}$$

with

$$\bar{\Phi}(t,x,\underline{x})=\bar{\Phi}^{(1)}(t,x,\underline{x})+\bar{\Phi}^{(2)}(t,x,\underline{x})+\frac{1}{2}\bar{\Phi}^{(3)}(t,x,\underline{x})\quad,$$

where, using that $\partial_t\mathcal{A}=\mathcal{A}\partial_t$ and again that $\partial_t u=-\mathcal{A}u$,

$$\begin{aligned}
&\bar{\Phi}^{(1)}(t,x,\underline{x})=(\partial_t+\bar{\mathcal{A}})(\partial_t u)(t,x,\underline{x})=-\bar{\mathcal{A}}(\mathcal{A}u)(t,x,\underline{x})+\mathcal{A}(\mathcal{A}u)(t,x)\\
&\bar{\Phi}^{(2)}(t,x,\underline{x})=\langle\bar{\mathcal{A}}(\partial_x u)(t,x,\underline{x})-\partial_x(\mathcal{A}u)(t,x),b(\underline{x})\rangle\\
\text{and}\qquad&\bar{\Phi}^{(3)}(t,x,\underline{x})=\mathrm{Tr}\left(\sigma^*(\underline{x})\left(\bar{\mathcal{A}}(\partial^2_{xx}u)(t,x,\underline{x})-\partial^2_{xx}(\mathcal{A}u)(t,x)\right)\sigma(\underline{x})\right).
\end{aligned}$$

By (2.10) and the sublinear growth of b and σ and the boundedness of its derivatives, one checks that there exists a positive real number r such that

$$|\bar{\Phi}(t,x,\underline{x})|\leq C_T(1+|x|^r+|\underline{x}|^r),\quad t\in[0,T].$$

Then, by Lemma 3.2(i) of [18],

$$\sup_{t\in[0,T]}\mathbb{E}[\mathbb{E}_{t_{k-1}}|\bar{\Phi}(t,\xi^{x,\mathbf{h}}_t,\xi^{x,\mathbf{h}}_{\underline{t}})|]\leq C_T\sup_{t\in[0,T]}\mathbb{E}[1+|\xi^{x,\mathbf{h}}_t|^r]\leq C_T(1+|x|^r).$$

Thus, it follows from (4.21)

$$|\mathbb{E}[g(X_T^x)] - \mathbb{E}[g(\xi_T^{x,\mathbf{h}})]| \leq C_T(1+|x|^r)\sum_{k=1}^{k_T}\int_{t_{k-1}}^{t_k}\int_{t_{k-1}}^{s} dvds = C_T(1+|x|^r)\sum_{k=1}^{k_T} h_k^2.$$

(ii) $\triangleright$ Assume first that $m=5$. Then, the function $x \mapsto u(t,x)$ is a $\mathcal{C}^5$-function whose existing partial derivatives have polynomial growth. One also observes that

$$\bar{\Phi}(v,\underline{x},\underline{x}) = \Phi_g(v,\underline{x})$$

where Φ_g is defined by (3.12). Thus, one rewrites (4.21) using (4.23) as follows

$$\mathbb{E}[g(X_T^x)] - \mathbb{E}[g(\xi_T^{x,\mathbf{h}})] = \frac{1}{2}\sum_{k=1}^{k_T}(h_k)^2\mathbb{E}[\Phi_g(t_{k-1},\xi_{t_{k-1}}^{x,\mathbf{h}})] + R_T(x,\mathbf{h})$$

where

$$R_T(x,h) = \sum_{k=1}^{k_T}\int_{t_{k-1}}^{t_k}\int_{t_{k-1}}^{s}\mathbb{E}_{t_{k-1}}\left(\bar{\Phi}(v,\xi_v^{x,\mathbf{h}},\xi_{\underline{v}}^{x,\mathbf{h}}) - \Phi_g(v,\xi_{\underline{v}}^{x,\mathbf{h}})\right)dvds.$$

Using that for any $\mathcal{C}^2$-function f and for every $x,\underline{x}\in\mathbb{R}^d$

$$|\bar{\mathcal{A}}f(x,\underline{x}) - \mathcal{A}f(x)| \leq C\left(|\nabla f(x)|.|\underline{x}-x| + \|D^2f(x)\|.|\underline{x}-x|.(1+|x|+|\underline{x}|)\right),$$

we deduce from the polynomial growth of the derivatives of u and the boundedness of the derivatives of b and σ that there exists $r, C_T>0$ such that for every $x,\underline{x}\in\mathbb{R}^d$,

$$|\bar{\Phi}(t,x,\underline{x}) - \Phi_g(t,\underline{x})| \leq C_T(1+|\underline{x}|^r+|x|^r)|x-\underline{x}|, \quad t\in[0,T].$$

By Lemma 3.2(i) and (ii) of [18] and the Hölder inequality, we obtain:

$$\mathbb{E}[|\bar{\Phi}(t,\xi_t^{x,\mathbf{h}},\xi_{\underline{t}}^{x,\mathbf{h}}) - \Phi_g(t,\xi_{\underline{t}}^{x,\mathbf{h}})|] \leq C_T\sqrt{t-\underline{t}}(1+|x|^{r'}), \quad t\in[0,T].$$

The announced control of $R(T,x)$ when $m=5$ follows.

$\triangleright$ Assume now that $m=6$. For every $i\geq 1$, we define recursively $\bar{\mathcal{A}}^{(i)}$ by $\bar{\mathcal{A}}^{(1)} = \bar{\mathcal{A}}$ and $\bar{\mathcal{A}}^{(i+1)}u(s,x,\underline{x}) = \bar{\mathcal{A}}(\bar{\mathcal{A}}^{(i)}u)(s,x,\underline{x})$. Then, by an Itô expansion of $(\bar{\Phi}(v,\xi_v^{x,\mathbf{h}},\xi_{\underline{v}}^{x,\mathbf{h}}))_{v\in[t_{k-1},t_k]}$ in (4.23), we obtain that

$$\mathbb{E}_{t_{k-1}}[u(t_k,\xi_{t_k}^{x,\mathbf{h}}) - u(t_{k-1},\xi_{t_{k-1}}^{x,\mathbf{h}})] \tag{4.24}$$

$$= \int_{t_{k-1}}^{t_k}\int_{t_{k-1}}^{s}\left(\Phi_g(t_{k-1},\xi_{t_{k-1}}^{x,\mathbf{h}}) + \int_{t_{k-1}}^{v}\mathbb{E}_{t_{k-1}}[(\partial_t+\bar{\mathcal{A}})\bar{\Phi}(w,\xi_w^{x,\mathbf{h}},\xi_{\underline{w}}^{x,\mathbf{h}})]dw\right)dvds$$

$$= \frac{(h_k)^2}{2}\Phi_g(t_{k-1},\xi_{t_{k-1}}^{x,\mathbf{h}}) + \rho_k \tag{4.25}$$

with

$$\rho_k = \int_{t_{k-1}}^{t_k}\int_{t_{k-1}}^{s}\int_{t_{k-1}}^{v}\mathbb{E}_{t_{k-1}}[\bar{\psi}(w,\xi_w^{x,\mathbf{h}},\xi_{\underline{w}}^{x,\mathbf{h}})]dwdvds$$

where $\bar{\psi}$ can be written

$$\bar{\psi}(t,x,\underline{x}) = \bar{\mathcal{A}}^{(3)}u(t,x,\underline{x}) - 3\bar{\mathcal{A}}^{(2)}\mathcal{A}u(t,x,\underline{x}),\underline{x}) + 3\bar{\mathcal{A}}\mathcal{A}^{(2)}u(t,x,\underline{x}) - \mathcal{A}^{(3)}u(t,x) \tag{4.26}$$

owing to the commutations between ∂_t and both operators $\bar{\mathcal{A}}$ and $\mathcal{A}$ respectively.
A precise computation shows that $\bar{\psi}$ is a polynomial function of $b(\underline{x})$, $\sigma(\underline{x})$ and of the derivatives of order up to 6 of b, σ and $x \mapsto u(t,x)$. Then, by Proposition 2.2, there exists $r > 0$ such that $|\bar{\psi}(t,x,\underline{x})| \leq C_T(1+|x|^r+|\underline{x}|^r)$. It follows that

$$|\rho_k| \leq C(h_k)^3(1+\mathbb{E}[|\xi^{x,\mathbf{h}}_{t_{k-1}}|^r]).$$

Summing (4.24) over k, we deduce that

$$\mathbb{E}[g(\xi^{x,\mathbf{h}}_T)] - \mathbb{E}[g(X^x_T)] = \mathbb{E}[u(T,\xi^{x,\mathbf{h}}_T) - u(0,x)] = \frac{1}{2}\sum_{k=1}^{k_T}(h_k)^2\mathbb{E}[\Phi_g(t_{k-1},\xi^{x,\mathbf{h}}_{t_{k-1}})] + R_T(x,\mathbf{h}).$$

Using that $\mathbb{E}[\sup_{t\in[0,T]}|\xi^{x,\mathbf{h}}_t|^r] \leq C(1+|x|^r)$, we derive

$$|R_T(x,\mathbf{h})| \leq C(1+|x|^r)\sum_{k=1}^{k_T}(h_k)^3 \leq C_T\|\mathbf{h}\|^2_\infty(1+|x|^r) \qquad \square$$

In the next lemma, we will denote by ψ_g the function defined by

$$\psi_g(s,\underline{x}) = \bar{\psi}(s,\underline{x},\underline{x}) \quad \forall (s,\underline{x}) \in [0,T]\times\mathbb{R}^d, \tag{4.27}$$

where $\bar{\psi}$ is defined by (4.26).

Lemma 4.2. *Assume the assumptions of Proposition 2.2 with $m = 7$. Then,*

$$\mathcal{E}_T(g,x,\mathbf{h}) = \sum_{k=1}^{k_T}\left(\frac{1}{2}(h_k)^2\mathbb{E}[\Phi_g(t_{k-1},\xi^{x,\mathbf{h}}_{t_{k-1}})] + \frac{1}{6}(h_k)^3\mathbb{E}[\psi_g(t_{k-1},\xi^{x,\mathbf{h}}_{t_{k-1}})]\right) + \widetilde{R}_T(x,\mathbf{h})$$

where Φ_g and ψ_g are defined by (3.12) *and* (4.27) *respectively and*

$$\widetilde{R}_T(x,\mathbf{h}) \leq C\|\mathbf{h}\|^{\frac{5}{2}}_\infty(1+|x|^r).$$

Proof. By (4.25),

$$\mathcal{E}_T(g,x,\mathbf{h}) = \sum_{k=1}^{k_T}\left(\frac{1}{2}(h_k)^2\mathbb{E}[\Phi_g(t_{k-1},\xi^{x,\mathbf{h}}_{t_{k-1}})] + \frac{1}{6}(h_k)^3\mathbb{E}[\psi_g(t_{k-1},\xi^{x,\mathbf{h}}_{t_{k-1}})] + \varepsilon_k\right)$$

with

$$\varepsilon_k = \int_{t_{k-1}}^{t_k}\int_{t_{k-1}}^{s}\int_{t_{k-1}}^{v}\mathbb{E}[\bar{\psi}(w,\xi^{x,\mathbf{h}}_w,\xi^{x,\mathbf{h}}_{\underline{w}}) - \psi_g(w,\xi^{x,\mathbf{h}}_{\underline{w}})]dwdvds.$$

Under the assumptions of the lemma, ψ_g is a $\mathcal{C}^1$-function such that $\nabla\psi_g$ has polynomial growth. As a consequence, there exists $r > 0$ such that

$$|\bar{\psi}(w,\xi^{x,\mathbf{h}}_w,\xi^{x,\mathbf{h}}_{\underline{w}}) - \psi_g(w,\xi^{x,\mathbf{h}}_{\underline{w}})| \leq C_T(1+|\underline{x}|^r+|x|^r)|x-\underline{x}|.$$

Then, we deduce from Hölder's inequality and Lemma 3.2(i) and (ii) of [18] that

$$|\varepsilon_k| \leq C(1+\mathbb{E}[|\xi^{x,\mathbf{h}}_{t_{k-1}}|^r])\int_{t_{k-1}}^{t_k}\int_{t_{k-1}}^{s}\int_{t_{k-1}}^{v}\sqrt{w-\underline{w}}\,dwdvds \leq C(1+|x|^r)h_k^{\frac{7}{2}}.$$

Summing over k completes the proof. $\square$

4.2 First and second order expansion of the weak error

In the next proposition, we apply Lemmas 4.1 and 4.2 to the sequence $\gamma^{(n)} := (\gamma_k^{(n)})_{k\geq 1}$ defined by

$$\gamma_1^{(n)} = \Gamma_{N(nT)+1} - nT \quad \text{and} \quad \gamma_k^{(n)} = \gamma_{N(nT)+k} \quad k \in \{2, \ldots, N((n+1)T) - N(nT) - 1\}$$

and $\gamma^{(n)}_{N((n+1)T)-N(nT)+1} = (n+1)T - \Gamma_{N((n+1)T)}$. In fact, $\gamma^{(n)}$ is well-defined if $\gamma_{N(nT)+1} < T$. This property is satisfied for large enough n.

Note that in the second part of the lemma, we use the notation Φ_{Φ_g}. We recall that, for a function $h : \mathbb{R}^d \to \mathbb{R}$, Φ_h is given by (3.12).

Proposition 4.5. *Let $m \in \mathbb{N}$ and $g : \mathbb{R}^d \to \mathbb{R}$ such that the assumptions of Proposition 2.2 hold. Then,*

(i)

$$\mathcal{E}_T(g, x, \gamma^{(n)}) = \frac{1}{2}\gamma_{N(nT)} \int_0^T \mathbb{E}[\Phi_g(s, X_s^x)]ds + r_n(x)$$

where Φ_g is defined by (3.12) *and*

$$|r_n(x)| \leq C_T(1 + |x|^r)) \begin{cases} \left(\gamma^{\frac{3}{2}}_{N(nT)} + (\gamma_{N(nT)} - \gamma_{N((n+1)T)})\right) & \text{if } m = 5 \\ \left(\gamma^{2}_{N(nT)} + (\gamma_{N(nT)} - \gamma_{N((n+1)T)})\right) & \text{if } m = 8. \end{cases} \tag{4.28}$$

(ii) Assume that $m = 9$ and that $\Gamma_{N(nT)} = nT$ for every $n \geq 1$. Then,

$$\mathcal{E}_T(g, x, \gamma^{(n)}) = \frac{1}{2}\gamma_{N(nT)} \int_0^T \mathbb{E}[\Phi_g(s, X_s^x)]ds + \gamma^2_{N(nT)} \int_0^T \mathbb{E}[\chi_g(s, X_s^x)]ds + \widetilde{\rho}_T(x, \gamma^{(n)}) \tag{4.29}$$

where

$$\chi_g(s, x) = \frac{1}{6}\psi_g(s, x) + \frac{1}{4}\Phi_{\Phi_g}(s, x) + \frac{1}{2}\left(\left(\mathcal{A} \circ \bar{\mathcal{A}}^{(2)} - \mathcal{A}^{(2)} \circ \bar{\mathcal{A}}\right) u_s\right)(x, x), \tag{4.30}$$

with $u_s(.) = u(s, .)$, ψ_g given by (4.27) *and*

$$|\widetilde{\rho}_T(x, \gamma^{(n)})| \leq C(1 + |x|^r)\left(\gamma_{N(nT)} - \gamma_{N((n+1)T)} + \gamma^{\frac{5}{2}}_{N(nT)}\right). \tag{4.31}$$

Proof. (*i*) For every $n \geq 0$, we want to use Lemma 4.1 with $\mathbf{h^{(n)}} := \gamma^{(n)}$. We denote by $t_k^{(n)}$ the discretization times related to $\mathbf{h^{(n)}}$ and by $k_T^{(n)}$ the index such that $t^{(n)}_{k_T^{(n)}} = T$. By the very definition of $\gamma^{(n)}$, we have $k_T^{(n)} = N((n+1)T) - N(nT) + 1$. Now, set $a_k^{(n)} = \mathbb{E}[\Phi_g(t^{(n)}_{k-1}, \xi^{x,\gamma^{(n)}}_{t^{(n)}_{k-1}})]$ and $S_k^{(n)} = \sum_{l=1}^k a_l^{(n)}\gamma_l^{(n)}$. By Lemma 4.1,

$$\mathbb{E}[g(\xi_T^{x,\gamma^{(n)}})] - \mathbb{E}[g(X_T^x)] = \sum_{k=1}^{k_T^{(n)}} (\gamma_k^{(n)})^2 a_k^{(n)} + R_T(x, \gamma^{(n)}) \text{ with}$$
$$\text{with} \quad |R_T(x, \gamma^{(n)})| \leq \begin{cases} C\gamma^{\frac{3}{2}}_{N(nT)}(1 + |x|^r) & \text{if } m = 5 \\ C\gamma^{2}_{N(nT)}(1 + |x|^r) & \text{if } m \geq 6, \end{cases}$$

where we used the fact that (γ_n) is non-increasing. In order to avoid the edge effects, we write

$$\mathbb{E}[g(\xi_T^{x,\gamma^{(n)}})] - \mathbb{E}[g(X_T^x)] = \sum_{k=2}^{k_T^{(n)}-1} (\gamma_k^{(n)})^2 a_k^{(n)} + r_n^{(0)}(x) + R_T(x,\gamma^{(n)})$$

with $r_n^{(0)}(x) = (\gamma_0^{(n)})^2 a_0^{(n)} + (\gamma_{k_T^{(n)}})^2 a_{k_T^{(n)}}^{(n)}$. Owing to the construction of $\gamma^{(n)}$ and the fact that $\Phi_g(x) \leq C_T(1+|x|^r)$ with $r > 0$, we have:

$$|r_n^{(0)}(x)| \leq C\gamma_{N(nT)}^2(1+|x|^r).$$

An Abel transform yields

$$\begin{aligned}
\sum_{k=2}^{k_T^{(n)}-1} (\gamma_k^{(n)})^2 a_k^{(n)} &= \gamma_{k_T^{(n)}-1} S_{k_T^{(n)}-1} - \gamma_2^{(n)} S_1^{(n)} - \sum_{k=2}^{k_T^{(n)}-2} (\gamma_{k+1}^{(n)} - \gamma_k^{(n)}) S_k^{(n)} \\
&= \gamma_{N((n+1)T)} \int_0^T \mathbb{E}[\Phi_g(s, X_s^x)]ds + r_n^{(1)}(x) + r_n^{(2)}(x)
\end{aligned}$$

where,

$$r_n^{(1)}(x) = \gamma_{N((n+1)T)} \int_0^T \mathbb{E}[\Phi_g(\underline{s}, \xi_{\underline{s}}^{x,\gamma^{(n)}}) - \Phi_g(s, X_s^x)]ds$$

and

$$\begin{aligned}
|r_n^{(2)}(x)| &\leq C\Big(\|\gamma^{(n)}\|_\infty^2 + \sum_{k=N(nT)+1}^{N((n+1)T)} (\gamma_k - \gamma_{k+1})\Big) \sup_{k\in\{1,\dots,k_T^{(n)}\}} |S_k^{(n)}| \\
&\leq C_T\Big(\|\gamma^{(n)}\|_\infty^2 + \sum_{k=N(nT)+1}^{N((n+1)T)} (\gamma_k - \gamma_{k+1})\Big) \sup_{s\in[0,T]} \mathbb{E}[|\Phi_g(\underline{s}, \xi_{\underline{s}}^{x,\gamma^{(n)}})|].
\end{aligned}$$

We derive from Lemma 3.2(i) of [18] and from the polynomial growth of Φ_g that there exists $r > 0$ such that

$$\sup_{s\in[0,T]} \mathbb{E}[|\Phi_g(\underline{s}, \xi_{\underline{s}}^{x,\gamma^{(n)}})|] \leq C(1+|x|^r)$$

which in turn implies

$$|r_n^{(2)}(x)| \leq C_T(1+|x|^r)\left((\gamma_{N(nT)} - \gamma_{N((n+1)T)}) + \gamma_{N(nT)}^2\right).$$

Now we focus on $r_n^{(1)}$ and we will inspect successively the cases $m = 5$ and $m = 8$.

$\triangleright$ *Case* $m = 5$: one derives from Proposition 2.2 and from the assumptions on the coefficients that $x \mapsto \Phi_g(x,t)$ is a $\mathcal{C}^1$-function such that $|\partial_x \Phi_g(x,t)| \leq C_T(1+|x|^r)$ where r is a positive real number. Thus,

$$\mathbb{E}[|\Phi_g(\underline{s}, X_s^x) - \Phi_g(\underline{s}, \xi_{\underline{s}}^{x,\gamma^{(n)}})|] \leq C_T \mathbb{E}[(1+|X_s^x|^r + |\xi_{\underline{s}}^{x,\gamma^{(n)}}|^r)|X_s^x - \xi_{\underline{s}}^{x,\gamma^{(n)}}|] \leq C_T\sqrt{s-\underline{s}}$$

owing to Lemma 3.2(i) and (ii) of [18] and to the Hölder inequality. Likewise, one checks that for every $i \in \{0,\dots,5\}$, there exists $r, C_T > 0$ such that for every $(s,t) \in [0,T]^2$ with $s \leq t$ and every $x \in \mathbb{R}^d$,

$$|\partial_{x^i}^i u(t,x) - \partial_{x^i}^i u(s,x)| \leq C_T(1+|x|^r)\sqrt{t-s}.$$

It follows that there exists $r>0$ such that, for every $t\in[0,T]$,

$$|\mathbb{E}[|\Phi_g(s,X_s^x)-\Phi_g(\underline{s},X_s^x)|]\leq C_T(1+|x|^r)\sqrt{s-\underline{s}}.$$

As a consequence,

$$\int_0^T \mathbb{E}[|\Phi_g(s,X_s^x)-\Phi_g(\underline{s},\xi_{\underline{s}}^{x,\gamma^{(n)}})|]ds\leq C_T\int_0^T\sqrt{s-\underline{s}}ds\leq C_T(1+|x|^r)\sqrt{\gamma_{N(nT)}},$$

and the announced control of r_n follows.

$\rhd$ *Case* $m=8$: We still need to focus on $r_n^{(1)}$. In this case, $x\mapsto\Phi_g(t,x)$ and $t\mapsto u(t,x)$ belong respectively to $\mathcal{C}^4(\mathbb{R}^d,\mathbb{R})$ and $\mathcal{C}^2([0,T],\mathbb{R})$. Then, on the one hand, using that $x\mapsto\partial_t\Phi_g+\mathcal{A}\Phi_g$ has polynomial growth, we deduce from an Itô expansion that,

$$\forall\, s\in[0,T],\quad |\mathbb{E}[\Phi_g(s,X_s^x)]-\mathbb{E}[\Phi_g(\underline{s},X_{\underline{s}}^x)]|\leq C_T(1+|x|^r)(s-\underline{s})$$

where r is a positive real number. On the other hand, for every (fixed) $s\in[0,T]$, we can apply Lemmma 4.1(i) to $\widetilde{g}$ defined by $\widetilde{g}(x)=\Phi_g(\underline{s},x)$ and $\mathbf{h}=\gamma^{(n)}$. We obtain that

$$\forall\, s\in[0,T],\quad |\mathbb{E}[\Phi_g(\underline{s},X_{\underline{s}}^x)]-\mathbb{E}[\Phi_g(\underline{s},\xi_{\underline{s}}^{x,\gamma^{(n)}})]|\leq C_{\underline{s}}(1+|x|^r)\sum_{k=1}^{k_{\underline{s}}^{(n)}}(\gamma_k^{(n)})^2\leq C_T(1+|x|^r)\gamma_{N(nT)}$$

since $t\mapsto C_t$ is non-decreasing. We finally derive the expected control for r_n

$$\int_0^T|\mathbb{E}[\Phi_g(s,X_s^x)]-\mathbb{E}[\Phi_g(\underline{s},\xi_{\underline{s}}^{x,\gamma^{(n)}})]|ds\leq C_T(1+|x|^r)\gamma_{N(nT)}.$$

(ii) Since for every $n\geq1$, $\Gamma_{N(nT)}=nT$, we have now $\gamma_k^{(n)}=\gamma_{N(nT)+k}$ for every $k\in\{1,\ldots,k_T^{(n)}\}$ where $k_T^{(n)}=N((n+1)T)-N(nT)$. Thus, $(\gamma^{(n)})$ is non-increasing and there is no "edge effect". Then, following the lines of (i) (with no edge effects) yields

$$\sum_{k=1}^{k_T^{(n)}}(\gamma_k^{(n)})^2\mathbb{E}[\Phi_g(t_{k-1}^{(n)},\xi_{t_{k-1}^{(n)}}^{x,\gamma^{(n)}})]=\gamma_{N((n+1)T)}\int_0^T\mathbb{E}[\Phi_g(\underline{s},\xi_{\underline{s}}^{x,\gamma^{(n)}})]ds+\widetilde{r}_n^{(1)}(x)$$

with
$$\widetilde{r}_n^{(1)}(x)=\sum_{k=1}^{k_T^{(n)}-1}(\gamma_k^{(n)}-\gamma_{k+1}^{(n)})\int_0^{t_k^{(n)}}\mathbb{E}[\Phi_g(\underline{s},\xi_{\underline{s}}^{x,\gamma^{(n)}})]ds$$

and

$$\sum_{k=1}^{k_T^{(n)}}(\gamma_k^{(n)})^3\mathbb{E}[\psi_g(t_{k-1}^{(n)},\xi_{t_{k-1}^{(n)}}^{x,\gamma^{(n)}})]=\gamma_{N((n+1)T)}^2\int_0^T\mathbb{E}[\psi_g(\underline{s},\xi_{\underline{s}}^{x,\gamma^{(n)}})]ds+\widetilde{r}_n^{(2)}(x)$$

with
$$\widetilde{r}_n^{(2)}(x)=\sum_{k=1}^{k_T^{(n)}-1}((\gamma_k^{(n)})^2-(\gamma_{k+1}^{(n)})^2)\int_0^{t_k^{(n)}}\mathbb{E}[\psi_g(\underline{s},\xi_{\underline{s}}^{x,\gamma^{(n)}})]ds.$$

It follows from Lemma 4.2 that

$$\mathcal{E}_T(g,x,\gamma^{(n)}) = \frac{\gamma_{N((n+1)T)}}{2}\int_0^T \mathbb{E}[\Phi_g(\underline{s},\xi_{\underline{s}}^{x,\gamma^{(n)}})]ds + \frac{\gamma^2_{N((n+1)T)}}{6}\int_0^T \mathbb{E}[\psi_g(\underline{s},\xi_{\underline{s}}^{x,\gamma^{(n)}})]ds \\ + \widetilde{r}_n^{(1)}(x) + \widetilde{r}_n^{(2)}(x) + \widetilde{R}_T(x,\gamma^{(n)}). \tag{4.32}$$

First, using that $(\gamma_k^{(n)})_k$ is non-increasing and that

$$\sup_{s\in[0,T]} \mathbb{E}[|\Phi_g(\underline{s},\xi_{\underline{s}}^{x,\gamma^{(n)}})| + |\psi_g(\underline{s},\xi_{\underline{s}}^{x,\gamma^{(n)}})|] \leq C_T(1+|x|^r)$$

with $r>0$, we deduce that

$$|r_n^{(1)}(x)| + |r_n^{(2)}(x)| \leq C(\gamma_{N(nT)+1} - \gamma_{N((n+1)T)})(1+|x|^r).$$

Second, let us control the right-hand side of (4.32). Applying the first claim (i) with $\widetilde{g} = \Phi_g$, we have for every $s\in[0,T]$,

$$\mathbb{E}[\Phi_g(\underline{s},\xi_{\underline{s}}^{x,\gamma^{(n)}}) - \Phi_g(\underline{s},X_{\underline{s}}^x)] = \frac{1}{2}\gamma_{N((n+1)T)}\int_0^{\underline{s}} \mathbb{E}[\Phi_{\Phi_g}(s,X_s^x)]ds + r_n(x)$$

with

$$|r_n(x)| \leq C_T\left(\gamma^{\frac{3}{2}}_{N(nT)} + (\gamma_{N(nT)+1} - \gamma_{N((n+1)T)})\right)(1+|x|^r) \tag{4.33}$$

since Φ_g is a $\mathcal{C}^5$-function whose derivatives of order $0,\ldots,5$ have polynomial growth under the assumptions of the lemma. We also deduce from Itô's formula that

$$\mathbb{E}[\Phi_g(\underline{s},X_{\underline{s}}^x) - \Phi_g(s,X_s^x)] = -\int_{\underline{s}}^s \mathbb{E}[(\partial_t + \mathcal{A})\Phi_g(s,X_s^x)]ds.$$

Then,

$$\int_0^T \mathbb{E}[\Phi_g(\underline{s},\xi_{\underline{s}}^{x,\gamma^{(n)}}) - \Phi_g(s,X_s^x)]ds = \frac{1}{2}\gamma_{N((n+1)T)}\int_0^T\int_0^s \mathbb{E}[\Phi_{\Phi_g}(s,X_s^x)]ds \\ + \gamma_{N((n+1)T)}\int_0^T \mathbb{E}\left[\left((\mathcal{A}\circ\bar{\mathcal{A}}^{(2)} - \mathcal{A}^{(2)}\circ\bar{\mathcal{A}})u_s\right)(X_s^x,X_s^x)\right]ds + \widetilde{r}_n^{(3)}(x)$$

with

$$|\widetilde{r}_n^{(3)}(x)| \leq C\gamma^2_{N(nT)+1}(1+|x|^r).$$

It follows that

$$\mathcal{E}_T(g,x,\gamma^{(n)}) - \frac{\gamma_{N((n+1)T)}}{2}\int_0^T \mathbb{E}[\Phi_g(s,X_s^x)]ds - \gamma^2_{N((n+1)T)}\int_0^T \mathbb{E}[\chi_g(s,X_s^x)]ds = \widetilde{\rho}_T(x,\gamma^{(n)})$$

with

$$\widetilde{\rho}_T(x,\gamma^{(n)}) = \widetilde{r}_n^{(1)}(x) + \widetilde{r}_n^{(2)}(x) + \widetilde{R}_T(x,\gamma^{(n)}) + \frac{\gamma_{N((n+1)T)}}{2}\widetilde{r}_n^{(3)}(x)$$

and the controls obtained previously yield (4.31). $\square$

5 Proof of Theorems 3.1 and 3.2

We introduce the notation $\mathbb{E}_{kT}$ to denote the conditional expectation $\mathbb{E}(\,.\,|\,\mathcal{F}_{kT})$. For every $n \geq 1$, $\bar{\mu}_n(\omega, f)$ can be decomposed as follows:

$$\bar{\mu}_n(\omega, f) - \nu(f) = \frac{1}{nT}\sum_{k=1}^{n} g(\xi_{(k-1)T}) - P_{\scriptscriptstyle T} g(\xi_{(k-1)T}) \tag{5.34}$$

$$= \frac{1}{nT}\left(g(\xi_0) - P_{\scriptscriptstyle T} g(\xi_{nT})\right) \tag{5.35}$$

$$+ \frac{1}{nT}\sum_{k=1}^{n-1} g(\xi_{kT}) - \mathbb{E}_{(k-1)T}[g(\xi_{kT})] \tag{5.36}$$

$$+ \frac{1}{nT}\sum_{k=1}^{n-1} \mathbb{E}_{(k-1)T}[g(\xi_{kT})] - P_{\scriptscriptstyle T} g(\xi_{(k-1)T}). \tag{5.37}$$

The proof of Theorems 3.1 and 3.2 is then based on the study of the rate of convergence of the three above terms.
First, (5.35) is a negligible term whose behavior is elucidated in Lemma 5.3 below. Second, we divide the study of the two main terms into two parts: Subsection 5.1 is devoted to the martingale component (5.36) for which we obtain some convergence in distribution results and in Subsection 5.2, we provide some convergence in probability results for (5.37) which strongly rely on the results of Section 5.2. Theorems 3.1 and 3.2 follow easily. Note that for Theorem 3.1, we only prove the case where f satisfies $(\mathcal{P}^{\mathrm{pol}}_{\mathbf{m},\mathbf{T}})$, the case where f satisfies $(\mathcal{P}^{\mathbf{b}}_{\mathbf{m},\mathbf{T}})$ can be handled likewise and is left to the reader.

Lemma 5.3. *Assume* $(\mathbf{S_{a,p}})$ *holds with* $p > 2$ *and* $a \in (0,1]$. *Assume that* $|g| \leq CV^r$ *for an* $r \in [0, \frac{p}{2})$. *Then,*

$$\sqrt{n}\left(\frac{1}{nT}\left(g(\xi_0) - P_{\scriptscriptstyle T} g(\xi_{nT})\right)\right) \xrightarrow{\mathbb{P}} 0 \quad \text{as } n \longrightarrow +\infty.$$

Proof. By Lemma 3 from [11], there exists $n_0 \in \mathbb{N}$ such that, for every $n \geq n_0$, there exist two positive real numbers $\widetilde{\beta}$ and $\widetilde{\alpha}$ such that

$$\mathbb{E}[V^p(\xi_{\Gamma_{n+1}})|\mathcal{F}_{\Gamma_n}] \leq V^p(\xi_{\Gamma_n}) + \gamma_{n+1} V^{p-1}(\xi_{\Gamma_n})\left(\widetilde{\beta} - \widetilde{\alpha} V^a(\xi_{\Gamma_n})\right).$$

Since $a > 0$ and $\lim_{|x|\to+\infty} V(x) = +\infty$, there exists $C_\alpha > 0$ such that $V^{p-1} \leq \frac{\widetilde{\alpha}}{2} V^{p+a-1} + C_\alpha$. It follows that there exists $\beta' > 0$ such that for every $n \geq n_0$,

$$\mathbb{E}[V^p(\xi_{\Gamma_{n+1}})] \leq \mathbb{E}[V^p(\xi_{\Gamma_n})] + \beta' \gamma_{n+1}.$$

Using that $\sup_{n \leq n_0} \mathbb{E}[V^p(\xi_{\Gamma_n})] < +\infty$ yields by induction the existence of a real constant $C > 0$ such that

$$\forall n \geq 1, \quad \mathbb{E}[V^p(\xi_{\Gamma_n})] \leq C\,\Gamma_n. \tag{5.38}$$

Thus, using Lemma 3.2(i) from [18], we obtain

$$|\mathbb{E}[P_{\scriptscriptstyle T} g(\xi_{nT})]| \leq C\mathbb{E}[V^r(\xi_{nT})] = O\big((\Gamma_{N(n,T)})^{\frac{r}{p}}\big) = O(n^{\frac{1}{2}-\varepsilon})$$

with $\varepsilon \in (0, 1/2)$ since $r/p < 1/2$. The result follows. $\square$

5.1 CLT for the martingale term

Proposition 5.6. *Assume* $(\mathbf{S_{a,p}})$ *holds with* $p>2$ *and* $a\in(0,1]$ *and that* $(\mathbf{S_T^\nu})$ *holds as well. Then, if* g *is a locally Lipschitz continuous function such that* $g\le CV^r$ *with* $2r<p/2+a-1$,

$$\frac{1}{\sqrt{nT}}\sum_{k=1}^{n} g(\xi_{kT})-\mathbb{E}_{(k-1)T}[g(\xi_{kT})] \overset{n\to+\infty}{\Longrightarrow} \mathcal{N}(0,\hat{\sigma}_g^2)$$

where $\hat{\sigma}_g^2=\dfrac{1}{T}\displaystyle\int g^2-(P_{_T}g)^2d\nu$.

Proof. Set $\zeta^{n,k}=\frac{1}{\sqrt{nT}}\left(g(\xi_{kT})-\mathbb{E}_{(k-1)T}[g(\xi_{kT})]\right)$. We aim at applying Lindeberg's Central Limit Theorem (see *e.g.* [5]) to the triangular array of $\mathcal{G}^n=(\mathcal{G}_k)_k$-martingale increments $\{\zeta^{n,k},k=1,\ldots,n,n\ge 1\}$. First, we have:

$$\sum_{k=1}^{n}\mathbb{E}[(\zeta^{n,k})^2/\mathcal{F}_{(k-1)T}]=\frac{1}{nT}\sum_{k=1}^{n}\Big(\mathbb{E}_{(k-1)T}[g^2(\xi_{kT})]-\left(\mathbb{E}_{(k-1)T}[g(\xi_{kT})]\right)^2\Big).$$

As a first step, we prove owing to a standard martingale argument that

$$\frac{1}{n}\sum_{k=1}^{n}\left(\mathbb{E}_{(k-1)T}[g^2(\xi_{kT})]-g^2(\xi_{kT})\right)\xrightarrow{n\to+\infty}0\quad a.s. \tag{5.39}$$

Indeed, denoting by $(N_n)_{n\ge1}$ the martingale defined by

$$N_n=\sum_{k=1}^{n}\frac{1}{k}\left(\mathbb{E}_{(k-1)T}[g^2(\xi_{kT})]-g^2(\xi_{kT})\right),$$

and using Lemma 3.2(i), Lemma 3.3(i) of [18] (with $\phi=V$ and $\delta_k=1/k^2$ respectively) and the fact that $2r\le\dfrac{p}{2}+a-1$, we obtain that *a.s.*,

$$\sum_{k=1}^{\infty}\mathbb{E}_{(k-1)T}[|\Delta N_k|^2]\le\sum_{k=1}^{\infty}\frac{1}{k^2}\mathbb{E}_{(k-1)T}[V^{\frac{p}{2}+a-1}(\xi_{kT})]\le C\sum_{k=1}^{\infty}\frac{1}{k^2}V^{\frac{p}{2}+a-1}(\xi_{(k-1)T})<+\infty.$$

Thus, (N_n) is an *a.s.* convergent martingale and the Kronecker Lemma yields (5.39).
Then, by Proposition 2.1(ii) and the fact that $g(x)=o(V^{\frac{p}{4}+\frac{a-1}{2}}(x))$ as $|x|\to+\infty$, $(\bar{\mu}_n(g^2))\xrightarrow{n\to+\infty}\nu(g^2)$ *a.s.* and it follows that

$$\frac{1}{n}\sum_{k=1}^{n}\mathbb{E}_{(k-1)T}[g^2(\xi_{kT})]\xrightarrow{n\to+\infty}\nu(g^2)\quad a.s.$$

On the other hand, let us show that

$$\frac{1}{n}\sum_{k=1}^{n}\left(\mathbb{E}_{(k-1)T}[g(\xi_{kT})]\right)^2\xrightarrow{n\to+\infty}\nu((P_{_T}g)^2). \tag{5.40}$$

First, by Lemma 3.2(i) and (ii) from [18], one can check that if h is a locally Lispchitz continuous function such that $h \leq CV^r$, then

$$P_{_T} h(x) - \mathbb{E}[h(\xi_T^{x,\mathbf{h^{(n)}}})] \xrightarrow{n\to+\infty} 0 \quad \text{locally uniformly in } x.$$

Then, using Proposition 2.1(i), we deduce that

$$\frac{1}{n}\sum_{k=1}^{n}\left(\mathbb{E}_{(k-1)T}[g(\xi_{kT})] - (P_{_T} g(\xi_{(k-1)T})\right)^2 \xrightarrow{n\to+\infty} 0 \quad a.s. \tag{5.41}$$

Now, since $(P_{_T} g)^2$ is continuous and dominated by V^{2r} with $2r < p/2 + a - 1$, (5.40) follows from Proposition 2.1(ii). We deduce that

$$\sum_{k=1}^{n} \mathbb{E}_{(k-1)T}[(\zeta^{n,k})^2] \xrightarrow{n\to+\infty} \hat{\sigma}_g^2 \quad a.s.$$

Finally, we have to check a Lindeberg-type condition $i.e.$ find $\delta > 0$

$$\sum_{k=1}^{n} \mathbb{E}_{(k-1)T}[|\zeta^{n,k}|^{2+\delta}] \xrightarrow{n\to+\infty} 0 \quad a.s.$$

Indeed, we choose δ such that $(2+\delta)r \leq p/2+a-1$ (keep in mind $2r < \frac{p}{2}+a-1$). We get

$$\begin{aligned}\sum_{k=1}^{n} \mathbb{E}_{(k-1)T}[|\zeta^{n,k}|^{2+\delta}] &\leq \frac{C}{n^{1+\frac{\delta}{2}}}\sum_{k=1}^{n}\mathbb{E}_{(k-1)T}[|g(\xi_{kT})|^{2+\delta}] \\ &\leq \frac{C_T}{n^{1+\frac{\delta}{2}}}\sum_{k=1}^{n} V^{(2+\delta)r}(\xi_{(k-1)T}) \leq \frac{C_T}{n^{\frac{\delta}{2}}}\sup_{n\geq 1}\bar{\mu}_n(V^{\frac{p}{2}+a-1}) \xrightarrow{n\to+\infty} 0\end{aligned}$$

owing to Proposition 2.1(i). □

The last term of the decomposition (5.37) is investigated in the next subsection.

5.2 Asymptotic behavior of the weak error term

In the next lemma, we show that tightness and convergence results for $(\bar{\mu}_n)$ can be extended to a wide class of weighted empirical measures related to the Euler scheme.

Lemma 5.4. *Assume that the conclusions of Proposition 2.1 hold true. Let $(a_n)_{n\geq 1}$ be a non-increasing sequence such that $A_n := \sum_{k=1}^n a_k \to +\infty$. Then,*

$$\sup_{n\geq 1}\frac{1}{A_n}\sum_{k=1}^{n} a_k V^{\frac{p}{2}+a-1}(\xi_{(k-1)T}) < +\infty \quad a.s.$$

and for every ν-a.s. continuous function f such that $f(x) = o(V^{\frac{p}{2}+a-1}(x))$ as $|x| \to +\infty$,

$$\frac{1}{A_n}\sum_{k=1}^{n} a_k f(\xi_{(k-1)T}) \to \nu(f).$$

Proof. By an Abel transform, for every function $h : \mathbb{R}^d \to \mathbb{R}$,

$$\frac{1}{A_n}\sum_{k=1}^{n} a_k h(\xi_{(k-1)T}) = \frac{na_n}{A_n}\bar{\mu}_n(\omega, h) + \frac{1}{A_n}\sum_{k=2}^{n}(k-1)(a_{k-1}-a_k)\bar{\mu}_{k-1}(\omega, h).$$

When $h(x) = V^{\frac{p}{2}+a-1}(x)$, we know by Proposition 2.1 that $M(\omega) := \sup_{n\geq 1}\bar{\mu}_n(\omega, V^{\frac{p}{2}+a-1}) < +\infty$ *a.s.* Then, as $(a_n)_{n\geq 1}$ is non-increasing, $na_n \leq A_n$ and we first deduce that

$$\sup_{n\geq 1}\frac{na_n}{A_n}\bar{\mu}_n(\omega, V^{\frac{p}{2}+a-1}) \leq M(\omega) \quad a.s.$$

Furthermore,

$$\frac{1}{A_n}\sum_{k=2}^{n}(k-1)(a_{k-1}-a_k)\bar{\mu}_{k-1}(\omega, V^{\frac{p}{2}+a-1}) \leq M(\omega)\frac{1}{A_n}\sum_{k=2}^{n}(k-1)(a_{k-1}-a_k).$$

Now, by another Abel transform, we have

$$0 \leq \frac{1}{A_n}\sum_{k=2}^{n}(k-1)(a_{k-1}-a_k) = \frac{a_1-(n-1)a_n}{A_n} + \frac{1}{A_n}\sum_{k=2}^{n-1}a_k \leq 2. \tag{5.42}$$

and the first assertion follows.

Let us now focus on the second part of the lemma. For every bounded continuous function f, set $h = f - \nu(f)$. By Proposition 2.1, $\bar{\mu}_n(\omega, h) \to 0$ as $n \to +\infty$. Then, following the lines of the first part of the proof, one easily gets that $\frac{1}{A_n}\sum_{k=1}^{n} a_k h(\xi_{(k-1)T}) \to 0$ *a.s.* as $n \to +\infty$. Since $\sup_{n\geq 1}\frac{1}{A_n}\sum_{k=1}^{n} a_k V^{\frac{p}{2}+a-1}(\xi_{(k-1)T}) < +\infty$ *a.s.*, this convergence can be extended to ν-a.s. continuous functions f satisfying $f(x) = o\big(V^{\frac{p}{2}+a-1}(x)\big)$ (as $|x| \to +\infty$) by standard uniform integrability arguments. $\square$

Proposition 5.7. *Assume* $(\mathbf{S_{a,p}})$ *holds with* $a \in (0,1]$ *and* $p = +\infty$. *Let* $m \in \mathbb{N}$ *and let* $g : \mathbb{R}^d \to \mathbb{R}$ *such that the assumptions of Proposition 2.2 hold.*

(i) If $m = 4$ *and* $\frac{1}{\sqrt{n}}\sum_{k=1}^{n}\gamma_{N(kT)} \xrightarrow{n\to+\infty} 0$,

$$\frac{1}{\sqrt{n}}\sum_{k=1}^{n}\Big|\mathbb{E}_{(k-1)T}[g(\xi_{kT})] - P_T g(\xi_{(k-1)T})\Big| \xrightarrow{n\to+\infty} 0 \quad a.s.$$

(ii) Assume $m = 5$ *and* $(\mathbf{S^{\nu}_T})$. *Then, if* $\sum_{k=1}^{+\infty}\gamma_{N(kT)} = +\infty$ *and if* $\gamma_{N(nT)} - \gamma_{N((n+1)T)} = o(\gamma_{N(nT)})$,

$$\frac{1}{\sum_{k=1}^{n}\gamma_{N(kT)}}\sum_{k=1}^{n}\mathbb{E}_{(k-1)T}[g(\xi_{kT})] - P_T g(\xi_{(k-1)T}) \xrightarrow{n\to+\infty} \frac{1}{2}\int\int_0^T \mathbb{E}[\Phi_g(t, X_t^x)]dt\,\nu(dx).$$

Proof. (i) Owing to Lemma 4.1 and the Markov property, there exists $r > 0$ such that

$$\begin{aligned}\sum_{k=1}^{n}\Big|\mathbb{E}_{(k-1)T}[g(\xi_{kT})] - P_T g(\xi_{(k-1)T})\Big| &\leq C_T\sum_{k=1}^{n}(1+|\xi_{(k-1)T}|^r)\left(\sum_{l=N((k-1)T)+1}^{N(kT)}\gamma_l^2\right)\\ &\leq C_T\sum_{k=1}^{n}\gamma_{N(kT)}(1+|\xi_{(k-1)T}|^r)\end{aligned}$$

since $(\gamma_n)_{n\geq 1}$ is nonincreasing and $\sum_{l=N((k-1)T)+1}^{N(kT)}\gamma_l \leq C_T$. First, since $\liminf_{|x|\to+\infty}\dfrac{V(x)}{|x|^\rho} > 0$ with $\rho > 0$, there exists a positive constant C such that $|x|^r \leq CV^{\frac{r}{\rho}}(x)$. If $\sum_{k=1}^{+\infty}\gamma_{N(kT)} < +\infty$, it follows from Lemma 3.3. from [18] applied with $\delta_k = \gamma_{N(kT)}$, that $\sum_{k=1}^{+\infty}\gamma_{N(kT)}(1+|\xi_{(k-1)T}|^r) < +\infty$. The result is then obvious in this case. Second, if $\sum_{k=1}^{+\infty}\gamma_{N(kT)} = +\infty$, Lemma 5.4 applied with $a_k = \gamma_{N(kT)}$ yields

$$\sup_{n\geq 1}\frac{1}{\sum_{k=1}^n\gamma_{N(kT)}}\sum_{k=1}^n\gamma_{N(kT)}(1+|\xi_{(k-1)T}|^r) < +\infty \quad a.s. \tag{5.43}$$

The result follows from the assumption $\dfrac{\sum_{k=1}^n\gamma_{N(kT)}}{\sqrt{n}} \xrightarrow{n\to+\infty} 0$.

(ii) Owing to Proposition 4.5(i), we have:

$$\begin{aligned}\frac{1}{\sum_{k=1}^n\gamma_{N(kT)}}\sum_{k=1}^n\mathbb{E}_{(k-1)T}[g(\xi_{kT})] - P_T g(\xi_{(k-1)T}) &= \frac{1}{\sum_{k=1}^n\gamma_{N(kT)}}\sum_{k=1}^n\gamma_{N(kT)}\varphi^{(1)}(\xi_{(k-1)T})\\ &\quad+\frac{1}{\sum_{k=1}^n\gamma_{N(kT)}}\sum_{k=1}^n r_n(\xi_{(k-1)T})\end{aligned}$$

where $\varphi^{(1)}(x) = \frac{1}{2}\int_0^T\mathbb{E}[\Phi_g(X_t^x)]dt$. Under our assumptions, $\varphi^{(1)}$ is continuous with polynomial growth. Then, by Lemma 5.4 applied with $a_k = \gamma_{N(kT)}$, we derive that

$$\frac{1}{\sum_{k=1}^n\gamma_{N(kT)}}\sum_{k=1}^n\gamma_{N(kT)}\varphi^{(1)}(\xi_{(k-1)T}) \xrightarrow{n\to+\infty} \int_{\mathbb{R}^d}\int_0^T\mathbb{E}[\Phi_g(X_s^x)]ds\nu(dx).$$

Finally, by (4.28) in Proposition 4.5 and the assumptions on the steps, $|r_n(x)| = o(\gamma_{N(nT)})(1+|x|^r)$ as $n\to+\infty$. Then, it follows from (5.43) that

$$\frac{1}{\sum_{k=1}^n\gamma_{N(kT)}}\sum_{k=1}^n r_n(\xi_{(k-1)T}) \xrightarrow{n\to+\infty} 0 \quad a.s. \qquad \square$$

5.3 Proof of Theorem 3.2

As mentioned previously, $(\mathbf{C_F}(\tfrac{1}{2}))$ is always true under the assumptions. Actually, F being Lipschitz continuous,

$$|\mathbb{E}[F(X^x) - F(\xi^{x,\mathbf{h}})]| \leq C\mathbb{E}[\sup_{t\in[0,T]}|X_t^x - \xi_t^{x,\mathbf{h}}|] \leq C(1+|x|)\sqrt{\|\mathbf{h}\|_\infty}$$

owing for instance to Lemma 3.2(ii) from [18]. As a consequence, we prove the two parts of Theorem 3.2 together under Assumption $(\mathbf{C_F}(\alpha))$ with $\alpha\in[1/2,1]$.
Setting $f_F(x) = \mathbb{E}[F(X^x)]$, $\bar{\mu}^{(n)}(\omega,F)$ can be written as follows:

$$\bar{\mu}^{(n)}(\omega,F) - \mathbb{E}_\nu(F) = \frac{1}{n}\sum_{k=1}^n\left(\Delta M_k^F + \mathcal{E}_T(F,\xi_{(k-1)T},\gamma^{(k)})\right) + \bar{\mu}_n(\omega,f_F) - \nu(f_F)$$

where $\Delta M_k^F = F(\xi^{((k-1)T)}) - \mathbb{E}_{(k-1)T}[F(\xi^{((k-1)T)})]$ and $\mathcal{E}_T(F, x, \mathbf{h}) = \mathbb{E}[F(\xi^{x,\mathbf{h}})] - \mathbb{E}[F(X^x)]$. Now, denoting by g_F the solution to (3.11) (when $f := f_F$), we can apply decomposition (5.34) to f_F. We obtain

$$\bar{\mu}^{(n)}(\omega, F) - \mathbb{E}_\nu(F) = \frac{1}{n}\sum_{k=1}^{n-1}\left(\Delta M_k^F + \frac{1}{T}(g_F(\xi_{kT}) - \mathbb{E}_{(k-1)T}[g_F(\xi_{kT})])\right) \tag{5.44}$$

$$+\frac{1}{n}\sum_{k=1}^{n}\mathcal{E}_T(F, \xi_{(k-1)T}, \gamma^{(k)}) + \frac{F(\xi^{(nT)}) - \mathbb{E}_{(n-1)T}[F(\xi^{(nT)})]}{n}$$

$$+\frac{1}{nT}(g_F(\xi_0) - P_T g_F(\xi_{nT})) + \frac{1}{nT}\sum_{k=1}^{n-1}\mathbb{E}_{(k-1)T}[g_F(\xi_{kT})] - P_T g_F(\xi_{(k-1)T}). \tag{5.45}$$

First, owing to the proof of Theorem 3.1, (5.45) is a negligible term under the step assumption. As well, under $(\mathbf{C_F}(\alpha))$,

$$\frac{1}{\sqrt{n}}\sum_{k=1}^{n}\left|\mathcal{E}_T(F, \xi_{(k-1)T}, \gamma^{(k)})\right| \le \frac{C}{\sqrt{n}}\sum_{k=1}^{n}(1 + |\xi_{(k-1)T}|)(\gamma_{N((k-1)T)})^\alpha.$$

Applying Lemma 5.4 with $a_k = \gamma_{(k-1)T}^\alpha$ and using that, under the assumptions of the theorem, $\frac{\sum_{k=1}^n(\gamma_{N((k-1)T)})^\alpha}{\sqrt{n}} \xrightarrow{n\to+\infty} 0$, we deduce that

$$\frac{1}{\sqrt{n}}\sum_{k=1}^{n}\left|\mathcal{E}_T(F, \xi_{(k-1)T}, \gamma^{(k)})\right| \xrightarrow{n\to+\infty} 0 \quad a.s.$$

Finally, F being Lipschitz continuous, $|F(\alpha)| \le C(1 + \sup_{t\in[0,T]}|\alpha(t)|)$ for every $\alpha \in \mathcal{C}([0,T], \mathbb{R}^d)$ and it follows from Lemma 3.2(i) from [18] that for every $n \ge 0$, $\mathbb{E}[|F(\xi^{(nT)}|] \le C(1 + \mathbb{E}[|\xi_{nT}|])$. Then, following the proof of Lemma 5.3, we obtain

$$\frac{F(\xi^{(nT)}) - \mathbb{E}_{(n-1)T}[F(\xi^{(nT)})]}{\sqrt{n}} \xrightarrow{\mathbb{P}} 0 \quad \text{as } n \longrightarrow +\infty.$$

Thus, it remains to control the right-hand member of (5.44). The proof is similar to that of Proposition 5.6. Set $\Delta\widetilde{M}_k = \Delta M_k^F + \frac{1}{T}(g_F(\xi_{kT}) - \mathbb{E}_{(k-1)T}[g_F(\xi_{kT})])$ and for $\alpha \in \mathcal{C}(\mathbb{R}_+, \mathbb{R}^d)$,

$$\kappa(\alpha) = F(\alpha) - \mathbb{E}[F(X^{\alpha(0)})] + \frac{1}{T}(g_F(\alpha(T)) - P_T g_F(\alpha(0))).$$

Denote by $X^{(kT),\xi_{kT}}$, the unique solution to $dY_t = b(Y_t)dt + \sigma(Y_t)dW^{(kT)}$ starting from ξ_{kT}. Using Lemma 3.2 from [18] and the properties of F and g_F, we obtain the existence of $r > 0$ such that

$$\mathbb{E}_{kT}[|\Delta\widetilde{M}_{k+1} - \kappa(X^{(kT),\xi_{kT}})|] \le C\sqrt{\gamma_{kT}}(1 + |\xi_{kT}|^r)$$

and such that

$$\mathbb{E}_{kT}[|\Delta\widetilde{M}_{k+1} + \kappa(X^{(kT),\xi_{kT}})|] \le C(1 + |\xi_{kT}|^r).$$

Owing to elementary inequality $a^2 - b^2 = (a-b)(a+b)$ and to the Schwarz inequality, we deduce

$$\frac{1}{n}\sum_{k=1}\mathbb{E}_{(k-1)T}[|\Delta\widetilde{M}_k|^2] - \mathbb{E}_{(k-1)T}[\kappa(X^{(kT),\xi_{kT}})^2] \xrightarrow{n\to+\infty} 0 \quad a.s.$$

Thus, applying Proposition 2.1(*ii*) to K defined by $K(x) = \mathbb{E}[\kappa(X^x)]$ (which is continuous with polynomial growth under the assumptions) yields

$$\frac{1}{n}\sum_{k=1}^{n}\mathbb{E}_{(k-1)T}[|\Delta\widetilde{M}_k|^2] \xrightarrow{n\to+\infty} \int K(x)\nu(dx) = \widetilde{\sigma}_F^2.$$

In order to conclude the proof, it remains to check Lindeberg's condition (see the proof of Proposition 5.6). This point is left to the reader.

6 Proof of Theorem 3.3

Like in (5.34), we can decompose $\widetilde{\mu}_n^{(\theta)}(\omega, f)$ as the sum of three terms:

$$\widetilde{\mu}_n^{(\theta)}(\omega,f) = \frac{1}{nT}\left(2g(\xi_0^{(\theta)}) - g(\xi_0) + 2P_T g(\xi_{nT}^{(\theta)}) - P_T g(\xi_{nT})\right)$$
$$+\frac{1}{nT}\sum_{k=1}^{n-1} 2\mathcal{E}_T(g,\xi_{(k-1)T}^{(\theta)},\widetilde{\gamma}^{(k-1)}) - \mathcal{E}_T(g,\xi_{(k-1)T},\gamma^{(k-1)}) + \frac{1}{nT}\sum_{k=1}^{n}(2\Delta M_k^{(2)} - \Delta M_k^{(1)})$$

where, for every sequence of positive real numbers $\mathbf{h} := (h_k)_{k=1}^{k_T}$, $\mathcal{E}_T(g,x,\mathbf{h})$ is defined by (4.20), for every $n \geq 1$, $\widetilde{\gamma}^{(n)}$ is defined by

$$\widetilde{\gamma}_1^{(n)} = \widetilde{\Gamma}_{\widetilde{N}(nT)+1} - nT, \qquad \widetilde{\gamma}_k^{(n)} = \widetilde{\gamma}_{\widetilde{N}(nT)+k},\ k\in\{2,\ldots,\widetilde{N}((n+1)T) - \widetilde{N}(nT) - 1\},$$

$\widetilde{\gamma}^{(n)}_{\widetilde{N}((n+1)T)-\widetilde{N}(nT)} = (n+1)T - \widetilde{\Gamma}_{\widetilde{N}((n+1)T)}$ and for every $n \geq 1$,

$$\Delta M_n^{(1)} = g(\xi_{nT}) - \mathbb{E}_{n-1}[g(\xi_{nT})] \quad \text{and} \quad \Delta M_n^{(2)} = g(\xi_{nT}^{(\theta)}) - \mathbb{E}_{n-1}[g(\xi_{nT}^{(\theta)})].$$

The method of proof is similar to that of Theorem 3.1. We study successively the three above terms: the first one is ruled by Lemma 5.3 which yields

$$\frac{1}{\sqrt{n}}\left(2g(\xi_0^{(\theta)}) - g(\xi_0) + 2P_T g(\xi_{nT}^{(\theta)}) - P_T g(\xi_{nT})\right) \xrightarrow{\mathbb{P}} 0 \quad \text{as } n \longrightarrow +\infty$$

under the assumptions of the theorem. Then, the asymptotic behavior of the second term is given by Lemma 6.6 below. Note that this is in this lemma that the "Romberg effect" comes out. To be precise, the Romberg extrapolation removes asymptotically the first order term of the weak error like in a constant step setting with finite horizon.

Finally, the asymptotic behavior of the third one is stated in Lemma 6.7.

Before stating Lemma 6.6, we need to extend the second part of Theorem 3.1 to some weighted empirical measures:

Lemma 6.5. *Assume* $(\mathbf{S}_{\mathbf{a},\infty})$ *holds with an* $a \in (0,1]$. *Assume* $(\mathbf{S}_{\mathbf{T}}^{\nu})$. *Let* $\alpha \in (0,1]$ *such that* b *and* σ *are* $\mathcal{C}^{5,\alpha}$*-functions on* $\mathbb{R}^d$ *with bounded derivatives. Let* $f : \mathbb{R}^d \to \mathbb{R}$ *be a Borel function satisfying* $(\mathcal{P}_{\mathbf{m},\mathbf{T}}^{\mathrm{pol}})$ *if* $p = +\infty$. *If moreover,* $(\eta_k)_{k\geq 1}$ *is a non-increasing sequence of positive numbers such that* $\sum_{k=1}^n \eta_k^2 \xrightarrow{n\to+\infty} +\infty$ *and* $\eta_n - \eta_{n+1} = o(\eta_n^2)$ *as* $n \to +\infty$, *then*

$$\frac{1}{\sum_{k=1}^n \eta_k^2}\left(\sum_{k=1}^n \eta_k f(\xi_{(k-1)T}) - \nu(f)\right) \xrightarrow{\mathbb{P}} \frac{1}{2T}\int_{\mathbb{R}^d}\int_0^T \mathbb{E}[\Phi_g(s,X_s^x)]ds\nu(dx)$$

as $n \longrightarrow +\infty$ *where* Φ_g *is defined by* (3.12).

Proof. We do not detail the proof of this result. The idea is to decompose $\sum_{k=1}^n \eta_k \big(f(\xi_{(k-1)T}) - \nu(f)\big)$ like in the beginning of Section 5. We then have a negligible term similar to (5.35), a martingale term $(\sum_{k=1}^n \eta_k^2)^{-1} \sum_{k=1}^n \eta_{k+1} \big(g(\xi_{kT}) - \mathbb{E}_{(k-1)T}[g(\xi_{kT})]\big)$ which goes *a.s.* to 0 (by a similar proof to that of (5.39), using Lemma 3.3 from [18] and the fact that $(\delta_n)_{n\geq 1}$ defined by $\delta_n = \eta_{n+1}^2(\sum_{k=1}^n \eta_k^2)^{-2}$ is a non-increasing sequence such that $\sum_{n\geq 1}\delta_n < +\infty$), a "weak error" term whose behavior is managed by Proposition 4.5 and Lemma 5.4 applied with $a_n = \eta_n$ and an additional term which is due to the fact that the sum is "weighted" and which is negligible under the assumption $\eta_n - \eta_{n+1} = o(\eta_n^2)$. □

Lemma 6.6. *Assume* $(\mathbf{S}_{\mathbf{a},\infty})$ *holds with an* $a \in (0,1]$. *Assume* $(\mathbf{S}_\nu^{\mathbf{T}})$. *Let* $g : \mathbb{R}^d \to \mathbb{R}$ *be such that the assumptions of Proposition 2.2 hold with* $m = 9$. *Assume that* $\varphi^{(1)}$ *defined by* $\varphi^{(1)}(x) = \frac{1}{2}\int_0^T \mathbb{E}[\Phi_g(s, X_s^x)]ds$ *satisfies* $(\mathcal{P}_{\mathbf{5},\mathbf{T}}^{\mathrm{pol}})$. *Assume that* $\sum_{k\geq 1}\gamma_{N(kT)}^2 = +\infty$ *and that* $\gamma_{(N(kt)+1} - \gamma_{N(kT)} = o(\gamma_{N(kT)}^2)$. *Then,*

(i) If $\frac{\sum_{k=1}^n \gamma_{N(kT)}^2}{\sqrt{n}} \xrightarrow{n\to+\infty} 0$,

$$\frac{1}{\sqrt{nT}}\sum_{k=1}^n \Big(2\,\mathcal{E}_T(g, \xi_{(k-1)T}^{(\theta)}, \widetilde{\gamma}^{(k-1)}) - \mathcal{E}_T(g, \xi_{(k-1)T}, \gamma^{(k-1)})\Big) \xrightarrow{\mathbb{P}} 0 \quad \text{as } n \longrightarrow +\infty.$$

(ii) If $(\gamma_n)_{n\geq 1}$ *is such that* $\Gamma_{N(kT)} = kT$ *for every* $n \geq 1$,

$$\frac{1}{\sum_{k=1}^n \gamma_{N(kT)}^2}\sum_{k=1}^n \Big(2\,\mathcal{E}_T(g, \xi_{(k-1)T}^{(\theta)}, \widetilde{\gamma}^{(k-1)}) - \mathcal{E}_T(g, \xi_{(k-1)T}, \gamma^{(k-1)})\Big) \xrightarrow{\mathbb{P}} \widetilde{m}_T T$$

as $n \longrightarrow +\infty$, *where* $\widetilde{m}_T$ *is defined in Theorem 3.3.*

Proof. The proofs of (i) and (ii) are similar being respectively based on the first and second order expansions obtained in Proposition 4.5. We choose to detail (ii). The crucial point is the fact that $\gamma_{N(kT)} = 2\widetilde{\gamma}_{\widetilde{N}(kT)}$ which allows us to write:

$$\begin{aligned} 2\mathcal{E}_T(g, \xi_{(k-1)T}^{(\theta)}, \widetilde{\gamma}^{(k-1)}) - \mathcal{E}_T(g, \xi_{(k-1)T}, \gamma^{(k-1)}) &= 2\left(\mathcal{E}_T(\xi_{(k-1)T}^{(\theta)}, \widetilde{\gamma}^{(k-1)}) - \widetilde{\gamma}_{\widetilde{N}(kT)}\nu(\varphi^{(1)})\right) \\ &- \left(\mathcal{E}_T(g, \xi_{(k-1)T}, \gamma^{(k-1)}) - \gamma_{N(kT)}\nu(\varphi^{(1)})\right). \end{aligned}$$

Owing to Proposition 4.5(ii),

$$\begin{aligned} \mathcal{E}_T(g, \xi_{(k-1)T}, \gamma^{(k-1)}) - \gamma_{N(kT)}\nu(\varphi^{(1)}) =& \gamma_{N(kT)}\left(\varphi^{(1)}(\xi_{(k-1)T}) - \nu(\varphi^{(1)})\right) \\ &+ \gamma_{N(kT)}^2\varphi^{(2)}(\xi_{(k-1)T}) + \widetilde{\rho}_T(\xi_{(k-1)T}, \gamma^{(n)}) \end{aligned}$$

with $\varphi^{(2)}(x) = \int_0^T \mathbb{E}[\chi_g(s, X_s^x)]ds$ and $|\widetilde{\rho}_T(\xi_{(k-1)T}, \gamma^{(n)})| \leq C\varepsilon(n)\gamma_{N(nT)}^2(1+|x|^r)$ where $\varepsilon(n) \to 0$ (since $\gamma_{N(nT)+1} - \gamma_{N((n+1)T)} = o(\gamma_{N(nT)}^2)$). Owing to Lemma 5.4 applied with $a_n = \gamma_{N(nT)}^2$, we obtain that

$$\frac{1}{\sum_{k=1}^n \gamma_{N(kT)}^2}\sum_{k=1}^n \widetilde{\rho}_T(\xi_{(k-1)T}, \gamma^{(k)}) \xrightarrow{n\to+\infty} 0 \quad a.s. \tag{6.46}$$

and that

$$\frac{1}{\sum_{k=1}^n \gamma_{N(kT)}^2}\sum_{k=1}^n \gamma_{N(kT)}^2\varphi^{(2)}(\xi_{(k-1)T}) \xrightarrow{n\to+\infty} \nu(\varphi^{(2)}) \quad a.s. \tag{6.47}$$

Finally, since $\varphi^{(1)}$ satisfies $(\mathcal{P}^{\text{pol}}_{\mathbf{5,T}})$, it follows from Lemma 6.5 applied with $\eta_k = \gamma_{N(kT)}$ that a.s.,

$$\frac{1}{\sum_{k=1}^n \gamma^2_{N(kT)}} \sum_{k=1}^n \gamma_{N(kT)} \left(\varphi^{(1)}(\xi_{(k-1)T}) - \nu(\varphi^{(1)})\right) \xrightarrow{n\to+\infty} \frac{1}{T}\int_{\mathbb{R}^d}\int_0^T \mathbb{E}[\Phi_{g_{\varphi^{(1)}}}(X^x_s)]ds\nu(dx) \qquad (6.48)$$

For the Euler scheme $(\xi^{(\theta)}_t)$, (6.46), (6.47) and (6.48) also hold replacing $\gamma_{N(kT)}$ by $\widetilde{\gamma}_{\tilde{N}(kT)}$. Then, the result follows by noticing that

$$\frac{2}{\sum_{k=1}^n \gamma^2_{N(kT)}} = \frac{1}{2\sum_{k=1}^n \widetilde{\gamma}^2_{\tilde{N}(kT)}}. \qquad \square$$

Lemma 6.7. *Assume* $(\mathbf{S}_{\mathbf{a},\infty})$ *holds with an* $a \in (0,1]$ *and assume that* $\mathrm{Tr}(\sigma\sigma^*(x)) = o(V^a(x))$ *as* $|x| \to +\infty$. *Let* $\theta \in [0,1]$ *and assume that* $Q^{(\theta)}_T$ *admits a unique invariant distribution* $\nu^{(\theta)}$. *Let* $g : \mathbb{R}^d \to \mathbb{R}$ *be a locally Lipschitz function with polynomial growth. Then,*

$$\frac{1}{\sqrt{nT}}\sum_{k=1}^n (2\Delta M^{(2)}_k - \Delta M^{(1)}_k) \overset{n\to+\infty}{\Longrightarrow} \mathcal{N}\left(0; (\hat{\sigma}^{(\theta)}_g)^2\right)$$

with $(\hat{\sigma}^{(\theta)}_g)^2 = 5\hat{\sigma}^2_g - \dfrac{4}{T}\displaystyle\int_{\mathbb{R}^d\times\mathbb{R}^d} (g(x)g(y) - P_T g(x) P_T g(y))\, \nu^{(\theta)}(dx,dy).$

Proof. The sketch of the proof is the same than that of Proposition 5.6 replacing $\zeta^{n,k}$ by $\widetilde{\zeta}^{n,k} = \frac{1}{\sqrt{nT}}(2\Delta M^{(2)}_k - \Delta M^{(1)}_k)$. First, we need to compute the limit of $\sum_{k=1}^n \mathbb{E}_{(k-1)T}[(\widetilde{\zeta}^{n,k})^2]$ as $n \to +\infty$:

$$\mathbb{E}_{(k-1)T}[(\widetilde{\zeta}^{n,k})^2] = \frac{4}{nT}\mathbb{E}_{(k-1)T}[(\Delta M^{(2)}_k)^2] - \frac{4}{nT}\mathbb{E}_{(k-1)T}[\Delta M^{(1)}_k \Delta M^{(2)}_k] + \frac{1}{nT}\mathbb{E}_{(k-1)T}[(\Delta M^{(1)}_k)^2].$$

By the proof of Proposition 5.6

$$\frac{1}{nT}\sum_{k=1}^n \mathbb{E}_{(k-1)T}[(\Delta M^{(i)}_k)^2] \xrightarrow{n\to+\infty} \hat{\sigma}^2_g \quad a.s.$$

Thus, it remains to deal with the “covariance” term:

$$\mathbb{E}_{(k-1)T}[\Delta M^{(1)}_k \Delta M^{(2)}_k] = \mathbb{E}_{(k-1)T}[g(\xi_{kT})g(\xi^{(\theta)}_{kT})] - \mathbb{E}_{(k-1)T}[g(\xi_{kT})]\mathbb{E}_{(k-1)T}[g(\xi^{(\theta)}_{kT})].$$

At this stage, we see that we need to apply Proposition 2.1 to $\mathbb{X}^{(\theta)} = (X, X^{(\theta)})$. In fact, using that $\mathrm{Tr}(\sigma\sigma^*) = o(V^a)$, one easily checks that $(\mathbf{S}_{\mathbf{a},\mathbf{p}})$ is satisfied for every $p > 0$ for $\mathbb{X}^{(\theta)}$ with the Lyapunov function $\widetilde{V}(x_1,x_2) = V(x_1) + V(x_2)$,

$$\widetilde{b}(x_1,x_2) = \begin{pmatrix} b(x_1) \\ b(x_2)\end{pmatrix} \quad \text{and} \quad \widetilde{\sigma}(x_1,x_2) = \begin{pmatrix} \sigma(x_1) & 0 \\ \theta\sigma(x_1) & \sqrt{1-\theta^2}\sigma(x_2)\end{pmatrix}.$$

Then, on the one hand, a martingale argument combined with Lemma 3.2(i) and (ii) of [18] yields

$$\frac{1}{n}\sum_{k=1}^n \left(\mathbb{E}_{(k-1)T}[g(\xi_{kT})g(\xi^{(\theta)}_{kT})] - g(\xi_{kT})g(\xi^{(\theta)}_{kT})\right) \xrightarrow{n\to+\infty} 0 \quad a.s.$$

On the other hand since $Q_T^{(\theta)}$ admits a unique invariant distribution $\nu^{(\theta)}$, we derive from Proposition 2.1(ii) that

$$\frac{1}{nT}\sum_{k=1}^{n}\mathbb{E}_{(k-1)T}[g(\xi_{kT})g(\xi_{kT}^{(\theta)})] \xrightarrow{n\to+\infty} \frac{1}{T}\int_{\mathbb{R}^d\times\mathbb{R}^d} g(x)g(y)\nu^{(\theta)}(dx,dy) \quad a.s.$$

Finally, we focus on the second part of $\mathbb{E}_{(k-1)T}[\Delta M_k^{(1)}\Delta M_k^{(2)}]$:

$$\begin{aligned}\mathbb{E}_{(k-1)T}[g(\xi_{kT})]\mathbb{E}_{(k-1)T}[g(\xi_{kT}^{(\theta)})] - P_T g(\xi_{(k-1)T})P_T g(\xi_{(k-1)T}^{(\theta)}) &= \mathbb{E}_{(k-1)T}[g(\xi_{kT}^{(\theta)})]\mathcal{E}(g,\xi_{(k-1)T},\gamma^{(n)})\\ &\quad - P_T g(\xi_{(k-1)T}^{(\theta)}\mathcal{E}(g,\xi_{(k-1)T}^{(\theta)},\widetilde{\gamma}^{(n)}).\end{aligned}$$

Thus, using that $\liminf_{|x|\to+\infty} V(x)/|x|^\rho > 0$ with $\rho > 0$, and following the arguments used to establish (5.41), we deduce that

$$\frac{1}{n}\sum_{k=1}^{n}\Big(\mathbb{E}_{(k-1)T}[g(\xi_{kT})]\mathbb{E}_{(k-1)T}[g(\xi_{kT}^{(\theta)})] - P_T g(\xi_{(k-1)T})P_T g(\xi_{(k-1)T}^{(\theta)})\Big) \xrightarrow{n\to+\infty} 0 \quad a.s.$$

Then, Proposition 2.1(ii) yields

$$\frac{1}{nT}\sum_{k=1}^{n}\mathbb{E}_{(k-1)T}[g(\xi_{kT})]\mathbb{E}_{(k-1)T}[g(\xi_{kT}^{(\theta)})] \xrightarrow{n\to+\infty} \frac{1}{T}\int P_T g(x)P_T g(y)\nu^{(\theta)}(dx,dy)$$

and we can conclude that $\sum_{k=1}^n \mathbb{E}_{(k-1)T}[(\widetilde{\zeta}^{n,k})^2] \xrightarrow{n\to+\infty} (\hat{\sigma}_g^{(\theta)})^2$ $a.s.$ Checking the Lindeberg condition like in the proof of Lemma 5.6 yields the announced result. $\square$

7 Asymptotic confluence and Poisson equation

For a symmetric real matrix A, denote by $\bar{\lambda}_A := \max\{\lambda_1,\dots,\lambda_d\}$ where $\lambda_1,\dots,\lambda_d$ denote the eigenvalues of A. For every $x\in\mathbb{R}^d$, we also denote by $A_\sigma(x)$ and $B_\sigma(x)$ the $d\times d$ real matrices defined by:

$$A_\sigma(x) := \sum_{k=1}^{r}\sum_{i=1}^{d}(\nabla\sigma_{i,k})^*\nabla\sigma_{i,k}(x) \quad \text{and} \quad B_\sigma(x) = \sum_{k=1}^{r}(\nabla\sigma^k)^*\nabla\sigma^k$$

with $\nabla\sigma^k = (\nabla\sigma_{1,k},\dots,\nabla\sigma_{d,k})$. We introduce the following assumption:

$(\mathbf{AC})_p$: For every $x\in\mathbb{R}^d$, $\nabla b(x) + \frac{1}{2}A_\sigma(x) + (p-1)B_\sigma$ is a negative definite matrix and

$$\sup_{x\in\mathbb{R}^d}\bar{\lambda}_{\nabla b+\frac{1}{2}A_\sigma+(p-1)B_\sigma} := -c_p < 0.$$

Remark 7.7. Note that when $d=1$, Assumption $(\mathbf{AC})_p$ is nothing but

$$\sup_{x\in\mathbb{R}}\Big(b'(x) + (p-\frac{1}{2})(\sigma'(x))^2\Big) < 0.$$

Proposition 7.8. *(i) Let k be a positive integer. Let $f : \mathbb{R}^d \to \mathbb{R}$ be a $\mathcal{C}^k$-function with bounded existing partial derivatives. Assume that b and σ are $\mathcal{C}^{n,\rho}$-functions with $\rho \in (0,1)$ and bounded derivatives of any order in $\{1,\ldots,n\}$. Let $T > 0$ and assume $(\mathbf{S^{\nu}_T})$. Assume $(\mathbf{AC})_{\mathbf{2k}}$. Then,*
(i) The function g_T defined by

$$\forall x \in \mathbb{R}^d, \quad g_T(x) = \sum_{n\geq 0} T(P_{nT}f(x) - \nu(f))$$

is a $\mathcal{C}^k$-function on $\mathbb{R}^d$ with polynomial growth as well as its existing partial derivatives which is solution to (3.11).
(ii) The function g_0 defined by

$$\forall x \in \mathbb{R}^d, \quad g_0(x) = \int_0^{\infty} (P_s f(x) - \nu(f))ds$$

is a $\mathcal{C}^k$-function on $\mathbb{R}^d$ with polynomial growth as well as its existing partial derivatives which is solution to the equation $\mathcal{A}g_0 = -(f - \nu(f))$.

The proof of this result is essentially based on Lemma 7.8 below.

Lemma 7.8. *(i) Assume that there exists $\rho \in (0,1)$ such that b and σ are $\mathcal{C}^{1,\rho}$-functions with bounded derivatives. Assume $(\mathbf{AC})_p$ with $p \geq 1$. Then, for every $x \in \mathbb{R}^d$, the tangent process (∇X_t^x) defined by $(\nabla X_t^x)_{i,j} = \partial_{x_j} X_t^{x,j}$ satisfies for every $j \in \{1,\ldots,d\}$:*

$$\mathbb{E}[|(\nabla X_t^x)_{.,j}|^{2p}]^{\frac{1}{2p}} \leq \exp(-c_p t) \quad \forall p \geq 2. \tag{7.49}$$

(ii) Let $n \in \mathbb{N}$. Assume that b and σ are $\mathcal{C}^{n,\rho}$-functions with $\rho \in (0,1)$ and bounded existing derivatives. Let $p \geq 1$ such that $(\mathbf{AC})_{np}$ holds. Then, there exist $C_{n,p} > 0$ and $\lambda_{n,p} > 0$ such that for every $x \in \mathbb{R}^d$,

$$\mathbb{E}[\|\nabla^{(n)}(X_t^x)\|^{2p}]^{\frac{1}{2p}} \leq C_{n,p}\exp(-\lambda_{n,p} t), \tag{7.50}$$

where for every $i_1,\ldots,i_n, j \in \{1,\ldots,d\}$, $(\nabla^{(n)} X_t^x))^j_{i_1,\ldots,i_n} = \partial^{(n)}_{x_{i_1},\ldots,x_{i_n}} X_t^{x,j}$ and $\|.\|$ denotes a norm on $\mathbb{R}^{d^{n+1}}$.

Proof. Set $Y_t^{i,j} = (\nabla X_t^x)_{i,j}$. For every $i,j \in \{1,\ldots,d\}$, $Y_0^{i,j} = \delta_{i,j}$ and

$$dY_t^{i,j} = \sum_{l=1}^d \partial_{x_l} b_i(X_t) Y_t^{l,j} dt + \sum_{k=1}^r \sum_{l=1}^d \partial_{x_l}\sigma_{i,k}(X_t) Y_t^{l,j} dB_t^k.$$

By the Itô formula, it follows that

$$(Y_t^{i,j})^2 = \delta_{i,j} + 2\sum_{l=1}^d \int_0^t Y_s^{i,j}\partial_{x_l} b_i(X_s) Y_s^{l,j} ds + \sum_{k=1}^r \int_0^t \left(\sum_{l=1}^d \partial_{x_l}\sigma_{i,k}(X_s) Y_s^{l,j}\right)^2 ds + M_t^{i,j}$$

where

$$M_t^{i,j} = 2\sum_{k=1}^r \int_0^t \left(\sum_{l=1}^d Y_s^{i,j}\partial_{x_l}\sigma_{i,k}(X_s) Y_s^{l,j}\right) dB_s^k.$$

Setting $Y^{\cdot,j}=(Y^{1,j},\ldots,Y^{d,j})^*$, we deduce that

$$|Y_t^{\cdot,j}|^2=1+2\int_0^t (Y_s^{\cdot,j})^*(\nabla b(X_s)+\frac{1}{2}A_\sigma(X_s))Y_s^{\cdot,j}ds+M_t^j$$

with $M_t^j=\sum_{i=1}^d M_t^{i,j}$. Let $\varepsilon>0$ and $q\geq 1$. Applying again the Itô formula with $f(x)=(\varepsilon+x)^q$ (with $x\geq 0$), we obtain that

$$(\varepsilon+|Y_t^{\cdot,j}|^2)^q=(\varepsilon+1)^q+2q\int_0^t(\varepsilon+|Y_s^{\cdot,j}|^2)^{q-1}(Y_s^{\cdot,j})^*(\nabla b(X_s)+\frac{1}{2}A_\sigma(X_s))Y_s^{\cdot,j}ds$$
$$+q\int_0^t(\varepsilon+|Y_s^{\cdot,j}|^2)^{q-1}dM_s^j+2q(q-1)\sum_{k=1}^r\int_0^t(\varepsilon+|Y_s^{\cdot,j}|^2)^{q-2}\left(\sum_{i,l=1}^d Y_s^{i,j}\partial_{x_l}\sigma_{i,k}(X_s)Y_s^{l,j}\right)^2 ds$$

Now,

$$\left(\sum_{i,l=1}^d Y_s^{i,j}\partial_{x_l}\sigma_{i,k}(X_s)Y_s^{l,j}\right)^2=|Y_s^{\cdot,j}|^2(Y^{\cdot,j})^*(\nabla\sigma^k)^*\nabla\sigma^k Y^{\cdot,j}.$$

Then, using that $(\nabla\sigma^k)^*\nabla\sigma^k$ is a non-negative symmetric matrix, we obtain that

$$(\varepsilon+|Y_t^{\cdot,j}|^2)^q=(\varepsilon+1)^q+2q\int_0^t(\varepsilon+|Y_s^{\cdot,j}|^2)^{q-1}(Y_s^{\cdot,j})^*H(X_s)Y_s^{\cdot,j}ds \tag{7.51}$$
$$+\widetilde{M}_t^{\varepsilon,j}$$

where $H=\nabla b+\frac{1}{2}A_\sigma+(q-1)B_\sigma$ and $(\widetilde{M}_t^{\varepsilon,j})_{t\geq 0}$ is a local martingale. Localizing the martingale and using that the derivatives of b and σ are bounded, it follows from the Gronwall lemma that for every $T>0$,

$$\mathbb{E}[(\varepsilon+|Y_t^{\cdot,j}|^2)^q]\leq C_{T,q}. \tag{7.52}$$

Thus, $(\widetilde{M}_t^{\varepsilon,j})$ is a (true) martingale. Set $q=p$. Using Assumption $(\mathbf{AC})_p$, we deduce from (7.51) that for every $\varepsilon>0$, for every $t\geq 0$.

$$\mathbb{E}[e^{2pc_pt}(\varepsilon+|Y_t^{\cdot,j}|^2)^p]\leq(\varepsilon+1)^p.$$

The result follows from Fatou's lemma.

(ii) We choose to write the sequel of the proof in the one-dimensional case. The generalization to the multi-dimensional is a direct adaptation of the following one but generates some tedious notations. First, by Theorem 3.3 p.223 of [8] and the conditions on the coefficients b and σ, $(t,x)\mapsto X_t^x$ is a $C^{n,\rho'}$-stochastic flow for every $\rho'\in(0,\rho)$. With a slight abuse of notation, we set $Y_t^{(k)}:=\partial_x^{(k)}X_t^x$. For every $k\in\{2,\ldots,n\}$ and every $x\in\mathbb{R}^d$, we can deduce from an induction that $Y_0^{(k)}=0$ and that

$$dY_t^{(k)}=\left(b'(X_t^x)Y_t^{(k)}+\sum_{l=2}^k b^{(l)}(X_t)P_{l,k}(Y_t^{(1)},\ldots,Y_t^{(k-l+1)})\right)dt$$
$$+\left(\sigma'(X_t^x)Y_t^{(k)}+\sum_{l=2}^k \sigma^{(l)}(X_t)P_{l,k}(Y_t^{(1)},\ldots,Y_t^{(k-l+1)})\right)dB_t$$

where for every $k \in \{2,\dots,n\}$, for every $l \in \{2,\dots,k\}$, $P_{l,k}$ is an homogeneous polynomial function with degree l. Applying Itô's formula, we deduce that for every $q \geq 1$,

$$|Y_t^{(k)}|^{2q} = 2q\int_0^t \left(b'(X_t^x)|Y_s^{(k)}|^{2q} + \sum_{l=2}^k b^{(l)}(X_s)P_{l,k}(Y_s^{(1)},\dots,Y_s^{(k-l+1)})\mathrm{sgn}(Y_s^{(k)})|Y_s^{(k)}|^{q-1}\right)dt \tag{7.53}$$
$$+\frac{2q(2q-1)}{2}\int_0^t |Y_s^{(k)}|^{2q-2}\left(\sigma'(X_s^x)Y_s^{(k)} + \sum_{l=2}^k \sigma^{(l)}(X_s)P_{l,k}(Y_s^{(1)},\dots,Y_s^{(k-l+1)})\right)^2 ds + N_t$$

where (N_t) is a local martingale. First, by (7.52) and an induction procedure, we deduce from Gronwall's lemma and the boundedness of the derivatives that

$$\forall\, q\geq 1,\quad \forall\, T>0,\quad \forall\, l\in\{1,\dots,k\},\quad \forall\, t\in[0,T],\quad \mathbb{E}[|Y_t^{(k)}|^{2q}] \leq C_{T,q,n}. \tag{7.54}$$

This implies in particular that (N_t) is a martingale. Set $q = p$. We now want to prove (7.50) under $(\mathbf{AC})_{kp}$ by an induction on $k \in \{1,\dots,n\}$. First, by (7.49), the assertion is true for $k = 1$. Now, assume that for every $p \geq 1$, for every $l \in \{1,\dots,k-1\}$, (7.50) is true under $(\mathbf{AC})_{lp}$ and suppose that $(\mathbf{AC})_{np}$ holds. Then, since $(\mathbf{AC})_{np}$ implies $(\mathbf{AC})_p$, we deduce from (7.53), the boundedness of the derivatives of b and σ and Itô's formula that for every $\varepsilon \in (0, c_p)$,

$$e^{(c_p-\varepsilon)t}\mathbb{E}[|Y_t^{(k)}|^{2p}] \leq \int_0^t e^{(c_p-\varepsilon)s}\left(-\varepsilon\mathbb{E}[|Y_t^{(k)}|^{2p}] + c\sum_{l=2}^k \mathbb{E}[\varphi_{l,k}(Y_s^{(1)},\dots,Y_s^{(k)})]\right)ds \tag{7.55}$$

with $c > 0$ and

$$\varphi_{l,k}(y_1,\dots,y_{k-l+1},y_k) = |P_{l,k}(y_1,\dots,y_{k-l+1})|.|y_k|^{2p-1} + (P_{l,k}(y_1,\dots,y_{k-l+1}))^2.|y_k|^{2p-2}.$$

By the Young inequality, it follows that for every $\varepsilon > 0$, there exists $C_\varepsilon > 0$ such that

$$|\varphi_{l,k}(y_1,\dots,y_{k-l+1},y_k)| \leq \frac{\varepsilon}{ck}|y_k|^{2p} + C_\varepsilon|P_{l,k}(y_1,\dots,y_{k-l+1})|^{2p}.$$

Thus,

$$\mathbb{E}[|Y_t^{(k)}|^{2p}] \leq C_\varepsilon e^{-(c_p-\varepsilon)t}\int_0^t e^{(c_p-\varepsilon)s}\sum_{l=2}^k \mathbb{E}[|P_{l,k}(Y_s^{(1)},\dots,Y_s^{(k-l+1)})|^{2p}]ds.$$

Set $\mathcal{N}_{l,k} := \{(\alpha_1,\dots,\alpha_{k-l+1}) \in \{0,\dots,l\}^{k-l+1}, \alpha_1+\ldots+\alpha_{k-l+1} = l, \sum_{i=1}^{k-l+1} i\alpha_i = k\}\}$. By an induction, one can check that $P_{l,k}$ is a linear combination of monomial functions Q such that $Q(x_1,\dots,x_{k-l}) = x_1^{\alpha_1}\dots x_l^{\alpha_{k-l+1}}$ with $(\alpha_1,\dots,\alpha_{k-l}) \in \mathcal{N}_{l,k}$. By the Hölder inequality applied with $p_m = k/(m\alpha_m)$ for every $m \in 1,\dots,k-l+1$, we deduce that,

$$\begin{aligned}\mathbb{E}[|P_{l,k}(Y_s^{(1)},\dots,Y_s^{(k-l+1)})|^{2p}] &\leq C\sum_{(\alpha_1,\dots,\alpha_{k-l+1})\in\mathcal{N}_{l,k}} \mathbb{E}[\prod_{m=1}^{k-l+1}|Y_s^{(m)}|^{2\alpha_m p}]\\ &\leq C\sum_{(\alpha_1,\dots,\alpha_{k-l+1})\in\mathcal{N}_{l,k}}\prod_{m=1}^{k-l+1}\mathbb{E}[|Y_s^{(m)}|^{\frac{2kp}{m}}]^{\frac{m\alpha_m}{k}}.\end{aligned}$$

It follows from the induction assumption and $(\mathbf{AC})_{kp}$ (in fact, $(\mathbf{AC})_{m(kp/m)}$) that there exists $\bar{\lambda} > 0$ such that for every $l \in \{2, \dots, k\}$, for every $m \in \{1, \dots, k-l+1\}$,

$$\mathbb{E}[|Y_s^{(m)}|^{\frac{2kp}{m}}] \leq C \exp(-\bar{\lambda} t)$$

where $\bar{\lambda}$ denotes a positive real number. Plugging the preceding controls in (7.55) yields

$$\mathbb{E}[|Y_t^{(k)}|^{2p}] \leq C_\varepsilon e^{(c_p - \varepsilon)} \int_0^t e^{c_p - \varepsilon - \rho} ds \quad (\rho > 0)$$

and the result follows. $\square$

Proof of Proposition 7.8. The proofs of (i) and (ii) being essentially the same ones, we only prove (i). First, one checks that g_T is well-defined on $\mathbb{R}^d$. Indeed, for every $n \in \mathbb{N}$

$$\begin{aligned}|P_{nT} f(x) - \nu(f)| &= |\int (P_{nT} f(x) - P_{nT} f(y)) \nu(dy)| \\ &= C \|\nabla f\|_\infty \int \sup_{x \in \mathbb{R}^d} \mathbb{E}[\|\nabla(X_{nT}^x)\|].|y - x| \nu(dy).\end{aligned}$$

Owing to Lemma 7.50(i), we deduce that there exists $c > 0$ such that for every $x \in \mathbb{R}^d$,

$$|P_{nT} f(x) - \nu(f)| \leq C(|x| + \int |y| \nu(dy)) e^{-cnT}$$

and the fact that g_T is well-defined follows. Second, by construction, g_T is clearly a solution to (3.11). Then, let us focus on the smoothness of g_T. First,

$$|\nabla(P_{nT} f(x) - \nu(f)| = |\mathbb{E}[\nabla f(X_{nT}^x).\nabla X_{nT}^x]| \leq C.\mathbb{E}[\|\nabla X_{nT}^x\|].$$

Then, the controls obtained in Lemma 7.50(i) and the Lebesgue Theorem of derivability show that g_T is $\mathcal{C}^1$ on $\mathbb{R}^d$. Following the proof of the preceding lemma and using the fact that the derivatives of f are bounded, we obtain again $\nabla^{(k)}(f(X_{nT}^x)$ can be controlled by the polynomial functions $P_{l,k}$ defined in the preceding proof. Then, the end of the preceding proof and the controls obtained in Lemma 7.50(ii) show that $\sum_{n \geq 1} \mathbb{E}[\|\nabla^{(k)}(f(X_{nT}^x)\|]$ is uniformly convergent and we deduce that g_T is a $\mathcal{C}^k$-function.

A Proof of Proposition 2.2

Under the assumptions of the Proposition, we derive from Theorem 3.3 p.223 of [8] that $x \to X_T^{(t),x}$ is $\mathcal{C}^{2k}$ on $\mathbb{R}^d$. It follows that $x \to g(X_T^{(t),x})$ is also $\mathcal{C}^{2k}$ on $\mathbb{R}^d$. Then, following the beginning of the proof of Lemma 7.8*(ii)*, we derive from the boundedness of the derivatives and from the Gronwall Lemma that for every $t \in [0, T]$, for every $l \geq 2k$, for every $r > 0$, $\mathbb{E}[\|\nabla^{(l)} X_T^{(t),x}\|^r] \leq C_T < +\infty$. Likewise, for every $r > 0$, we also deduce from the Gronwall Lemma and the sublinear growth of b and σ that for every $t \in [0, T]$, $\mathbb{E}[|X_T^{(t),x}|^r] = \mathbb{E}[|X_{T-t}^x|^r] \leq C_T(1 + |x|^r)$. Now, for every $l \leq 2k$, $\nabla^{(l)}(g(X_T^{(t),x}))$ is a polynomial function of $\nabla^{(1)} X_T^{(t),x}$, ..., $\nabla^{(l)} X_T^{(t),x}$ and $\nabla g(X_T^{(t),x}), \dots, (\nabla^{(l)} g)(X_T^{(t),x})$. Thus, using that the derivatives of g have polynomial growth and the preceding controls, we deduce from the

Hölder inequality that for every $r > 0$, there exists $\rho > 0$ such that for every $l \in \{1, \dots, k\}$, for every $i_1, \dots, i_l \in \{1, \dots, d\}$, for every $x \in \mathbb{R}^d$,

$$\mathbb{E}[|\partial^l_{x_{i_1},\dots,x_{i_l}} \left(g(X_T^{(t),x})\right) |^r] \leq C_T(1 + |x|^\rho)$$

where C_T does not depend on x. The fact that $x \to u(t,x)$ is $\mathcal{C}^k$ on $\mathbb{R}^d$ then follows from a uniform integrability argument and the growth control can be directly obtained taking $r = 1$ in the preceding inequality. Finally, using that for a $\mathcal{C}^2$-function h with polynomial growth, $t \to \mathbb{E}[h(X^x_{T-t})]$ is $\mathcal{C}^1$ on $[0,T]$ with $\partial_t \mathbb{E}[h(X^x_{T-t})] = -\mathbb{E}[\mathcal{A}h(X^x_t)]$, we deduce from an iteration that $t \to u(t,x)$ is $\mathcal{C}^k$ on $\mathbb{R}^d$ since g, b and σ are $\mathcal{C}^{2k}$ on $\mathbb{R}^d$.

B Proof of (3.15)

We prove the result under $(\mathbf{AC})_2$ which ensures

- the representations of g_T and g_0 given in Proposition 7.8,
- the exponential convergence of the semi-group toward the invariant distribution at least for locally Lispchitz continuous functions with subquadratic growth.

Let f be a $\mathcal{C}^2$-function such that f, ∇f and $D^2 f$ are bounded. Without loss of generality, we also assume that $\nu(f) = 0$. First, using that

$$g_T(x) - g_0(x) = -\int_0^\infty (P_s - P_{\lfloor \frac{s}{T} \rfloor T}) f(x) ds,$$

it follows from the continuity of $t \mapsto P_t f(x)$ and from the exponential convergence of $P_t f$ to $\nu(f) = 0$ that $g_T(x) \to g_0(x)$ for every x. Then, since

$$\frac{1}{T} g_T^2 - (P_T g_T)^2 = (g_T + P_T g_T) \frac{g_T - P_T g_T}{T} = (g_T + P_T g_T) f,$$

it follows that $\hat{\sigma}^2_T = -2 \int g_0(x) \mathcal{A} g_0(x) \nu(dx) = \hat{\sigma}^2_0$ as $T \to 0$.
Second, using that $\partial P_{nT} f = n P_{nT} \mathcal{A} f$

$$\partial_T g_T(x) = \sum_{n \geq 0} \left(P_{nT} f(x) + nT P_{nT} \mathcal{A} f(x) \right).$$

On the one hand, since for every $n \in \mathbb{N}$, $0 = \nu(f) = P_{nT} f(x) + \int_{nT}^{+\infty} P_s(\mathcal{A}f)(x) ds$, it follows from a change of variable that

$$\partial_T g_T(x) = \sum_{n \geq 0} \int_{nT}^{+\infty} (P_{\lfloor \frac{s}{T} \rfloor T} \mathcal{A} f - P_s \mathcal{A} f)(x) ds.$$

The continuity of $t \mapsto P_t \mathcal{A} f$ combined with the exponential convergence of the semi-group yields $\partial_T g_T \xrightarrow{T \to 0} 0$.
On the other hand, using again the exponential convergence and the fact that $\nu(f) = \nu(\mathcal{A}f) = 0$, we have for every $T > 0$

$$|\partial_T g_T(x) - f(x)| \leq C \sum_{k \geq 1} e^{-\lambda_1 kT} + kT e^{-\lambda_2 kT} \leq C \left(\frac{e^{-\lambda_1 T}}{1 - e^{-\lambda_1 T}} + (1 + |x|) \frac{T e^{-\lambda_2 T}}{1 - e^{-\lambda_2 T}} \right)$$

where C, λ_1 and λ_2 are some positive numbers. As a consequence, $\partial_T g_T(x) - f(x) \to 0$ as $T \to +\infty$ and the dominated convergence theorem yields the last statement of (3.15).

References

[1] Basak G.K., Bhattacharya R.N. (1992). Stability in Distribution for a Class of Singular diffusions, *Ann. Probab.*, **20**(1):312–321.

[2] Bhattacharya R.N. (1982). On the functional Central Limit Theorem and the law of the iterated logarithm for Markov processes, *Z. Wahrsch. Verw. Gebiete.*, **60**(2):185–201.

[3] Billingsley P. (1968) *Convergence of Probability Measures*, Wiley. 02333096

[4] Bouleau N., Lépingle D. (1994). *Numerical methods for stochastic processes*, John Wiley & Sons, Inc., New York, 359 p. 1274043

[5] Hall P., Heyde C. (1980).*Martingale Limit Theory and its Application*, Academic Press.

[6] Has'minskii R.Z. (2012), *Stochastic Stability of Differential Equations*, Springer, 2nd Edition, 339 p.

[7] Karlin S., Taylor H.M., *A second course in stochastic processes*, Academic Pres, New York, 1981, 452p.

[8] Kunita H. (1982). Stochastic differential equations and stochastic flows of diffeomorphisms. *Cours d'école d'été de Saint-Flour*, LN-1097, Springer-Verlag, 1982.

[9] Ladyzhenskaya, O. A., Ural'tseva, N. N. (1968). *Linear and quasilinear elliptic equations.* Academic Press, New York-London:495 p. 244627

[10] Lamberton D. and Pagès G. (2002). Recursive computation of the invariant distribution of a diffusion. *Bernoulli* **8**:367-405. 1913112

[11] Lamberton D. and Pagès G. (2003). Recursive computation of the invariant distribution of a diffusion: The case of a weakly mean reverting drift. *Stoch. Dynamics* **4**:435-451. 2030742

[12] Lapeyre B. and Temam E. (2001). Competitive Monte Carlo methods for the pricing of Asian Options. *Journal of Computational Finance* **5**(1):39-59.

[13] Lemaire V. (2007). An adaptive scheme for the approximation of dissipative systems. *Stochastic Process. Appl.* **117**:1491-1518. 2353037

[14] Lemaire V. (2005). *Estimation numérique de la mesure invariante d'un processus de diffusion*, PhD Thesis, Université de Marne-La Vallée.

[15] Lemaire V., Pagès G. and Panloup F. (2013). Invariant measure of duplicated diffusions and application to Richardson-Romberg extrapolation, pre-print LPMA 1565, http://hal.archives-ouvertes.fr/hal-00785766.

[16] Pagès, G. (2001). Multi-step Richardson-Romberg extrapolation: remarks on variance control and complexity. *Monte Carlo Methods Appl.* **(13)**(1):37-70.

[17] Pagès G., Panloup F. (2009). Approximation of the distribution of a stationary Markov process with application to option pricing. *Bernoulli*, **15**(1):146-177. 2546802

[18] Pagès G., Panloup F. (2012). Ergodic approximation of the distribution of a stationary diffusion : Rate of convergence. *Ann. Applied Probab.*, **22**(3) (2012):1059-1100

[19] Panloup F. (2008). Recursive computation of the invariant measure of a SDE driven by a Lévy process. *Ann. Appl. Probab.* **18**:379-426. 2398761

[20] Panloup F. (2008). Computation of the invariant measure of a Lévy driven SDE: Rate of convergence, *Stochastic Process. Appl.*, **118**(8):1351–1384. 2427043

[21] Pardoux E., Veretennikov A. (2001). On Poisson equation and diffusion approximation I, *Ann. Probab.*, **29**(3):1061–1085. 1872736

[22] Pardoux E., Veretennikov A. (2003). On Poisson equation and diffusion appoximation II, *Ann. Probab*, **31**(3):1166–1192. 1988467

[23] Pardoux E., Veretennikov A. (2006). On the Poisson equation and diffusion approximation III, *Ann. Probab*, **33**(3):1111–1133. 2135314

[24] Talay D., Tubaro L. (1990). Expansion of the global error for numerical schemes solving stochastic differential equations, *Stoch. Anal. Appl.*, **8**:94–120.